

# On the road to REDD

An emissions trading scheme gives forests a market value on the basis of how much carbon they sequester. It could help to control global warming — if developing nations meet their responsibilities.

Environmentalists have spent decades working to protect tropical forests, both to promote biodiversity and to conserve nature's bounty. All too often those efforts have fallen short in the face of economic forces that put a higher price on timber and cleared land than on the forests themselves. But that may soon change if international climate negotiators can include forest carbon in a treaty to control global warming. The path forward will not be easy, as the News Feature on Madagascar's forests on page 26 points out, but it is surely a worthy experiment.

Forest carbon represents a new way of thinking about conservation, one that measures forests in terms of the carbon they sequester in their biomass and soil. The numbers are substantial: deforestation is currently responsible for up to 20% of global carbon emissions, which means that protecting forests could noticeably slow global warming. Doing so will be difficult, though, given the social issues at play and the weak governance in many tropical nations. But it should be relatively cheap compared with other methods of reducing carbon emissions. And with carbon footing the bill, tropical forests might finally get the kind of attention — and resources — they deserve.

As the world prepares for the United Nations climate summit in Copenhagen this December, negotiations over how a new climate treaty would incorporate a market for forest conservation credits — a trading system known as 'reducing emissions from deforestation and forest degradation' (REDD) — have been among the most fruitful to date. REDD negotiators might well be closer to a deal than any of their treaty counterparts working on emissions targets, financing and the like. Developed countries see REDD as a potentially cheap and beneficial way to reduce emissions, and developing countries see it as a cash infusion that could be used to promote a new model of sustainable development.

At present, REDD pilot projects are sprouting up in communities around the tropics, often using government funds or in some cases

carbon credits that have been issued on voluntary carbon markets. As helpful as these individual projects might be for improving people's livelihoods and preserving local biodiversity, however, it's not clear that they measurably reduce global-warming emissions. To realize the full promise of REDD — and to tap into the much larger flows of private money expected in future carbon markets — nations must ultimately manage their forests on a national scale. This means that they will need to beef up their science and regulatory infrastructure in order to inventory all their forest carbon, show that they can control land use at the local level and prove that their emissions are declining. Exactly how they achieve this will probably vary by country, and that is fine. As long as forests are left standing and emissions are going down, nations should have some flexibility to set up systems that most benefit their own people.

Many pitfalls lie ahead. As the situation in Madagascar shows, political instability can derail environmental reforms; continuing poverty and bad policy, as well as droughts and fires, could do the same. But governments, climate negotiators and environmentalists are working on solutions to these challenges, and there is no evidence yet that they cannot be overcome.

One thing is clear: it is at the local level that forest protection will either succeed or fail. Governments must find ways to address the social and economic problems that push people to cut down their forests, as well as instituting laws against doing so. And although it would be foolhardy to think of REDD as a panacea, the idea certainly dovetails nicely with development goals. Indeed, those nations that are able to spread the wealth in such a way that local communities see benefits are the most likely to succeed. ■

**"Governments must find ways to address the social and economic problems that push people to cut down their forests."**

## A drug-induced low

The sacking of a government adviser on drugs shows Britain's politicians can't cope with intelligent debate.

During his tenure as the UK government's chief adviser on drug abuse, David Nutt ruffled many a feather with his provocative remarks. Earlier this year, for example, he published an article that called for a wider debate on society's approach to risk and that favourably compared the dangers of the psychoactive amphetamine drug MDMA (ecstasy) to those of horse-riding (D. J. Nutt *J. Psychopharmacol.* 23, 3–5; 2009).

But it was only on 30 October that Nutt, a professor in neuro-

psychopharmacology at Imperial College London, was summarily fired from his position as chair of the British government's Advisory Council on the Misuse of Drugs by home secretary Alan Johnson. According to Johnson, Nutt's crime was to muddy the allegedly clear waters of government drugs policy by publicly making statements that questioned it, thereby going beyond his remit as a scientific adviser (see *Nature* doi:10.1038/news.2009.1053; 2009).

That concern should not be dismissed lightly. Politicians cannot always base their decisions solely on scientific advice, but must also consider such factors as public sentiment. Scientific advisers who publicly attack decisions they consider to be less than ideal, and in so doing provide ammunition for political opponents of those decisions, are entering dangerous territory.

Nonetheless, in this case, the position of the Labour government and

of the leading opposition party, the Conservatives, which vigorously supported Nutt's sacking, has no merit at all. It deals a significant blow both to the chances of an informed and reasoned debate over illegal drugs, and to the parties' own scientific credibility.

Nutt's fate was decided following the publication last week of a thoughtful analysis of the challenges in estimating the harm done by drugs (see [go.nature.com/dPiUAt](http://go.nature.com/dPiUAt)). In this paper, for example, Nutt notes that public perception is highly influenced by the way the media cover dramatic events such as drug deaths. He goes on to analyse such coverage over a decade. The data show that the media have disproportionately highlighted the comparatively small number of deaths caused by drugs such as cocaine and MDMA compared with the far more numerous deaths caused by other substances, such as alcohol.

The imbalance has convinced the public — and politicians — that some notorious drugs are much more dangerous than they are. Such perceptions heavily influence the government's classification of substances under the Misuse of Drugs Act, which in turn determines the penalty for being found in possession of a given drug.

Nutt goes on to detail efforts to develop an evidence-based scheme that ranks drugs according to “nine parameters of harm”, which range from an individual drug user's ravaged veins to society's extra

payments for health care. The harm ranking that results bears only an approximate resemblance to the official classification. For example, alcohol and tobacco, which don't fall under the Misuse of Drugs Act, rank considerably higher than MDMA, which is classified as among the most dangerous drugs.

That scheme has no official standing, unfortunately. But in describing it, Nutt attempted to do what he has done in many other public discussions: portray risk in terms that people can understand, look at the underlying factors that influence scientific and public debate, and thereby highlight how politicized discussions about drug use have become. To Johnson, this was apparently tantamount to “lobbying” against political decisions.

Thus the sacking. As *Nature* went to press, two members of Nutt's former committee had resigned in protest against the government's intolerance. They were right to do so. The government, meanwhile, badly needs to restore its credibility on this issue. One good way to do that would be to follow Nutt's suggestion to turn the advisory council into an independent body reporting to parliament as a whole, not to any individual official. An independent, scientifically run drug-regulation system would also free politicians from having to politick over who is toughest on drugs, something that would spare them and scientists much unnecessary future trouble. ■

## 140 years on

*Nature's* birthday offers an occasion to reflect on the past and look to the future.

**N**ature's first issue appeared on Thursday 4 November 1869. 7,269 issues later, a little bit of satisfaction may be in order given that the journal has survived wars (during which publication was suspended at least once) and, so far, the Internet's onslaught on traditional models of publishing. *Nature's* papers are highly cited for what seem to us to be good reasons. Lots of people (millions online every month) want to read the journal. So where do we need to be self-critical? Readers will no doubt have many answers, but here are a few.

One is that we need to keep an eye on, and even counter, the way in which *Nature* is used or abused by others.

Throughout its history, and certainly over the past few decades, *Nature's* goal as a journal has been to publish papers that make deep and broad impacts in and beyond the natural sciences. We have pursued that goal in a spirit of vigorous independence. Because *Nature's* editors are free from any association with societies or editorial boards or immediate commercial imperatives, we have always made the final call on what should be published. It is not unknown for us to override the negative recommendations of three referees about a technically valid paper because we believe it to be more interesting than any of them perceive it to be. The responses of the community and the citations generally seem to have validated our judgements.

But others sometimes put more weight on our judgement than it can justifiably bear. Large grants, philanthropic donations and personal chairs have been awarded on the strength of a paper in *Nature* — in effect, using editors' decisions as a surrogate for

independent judgement. This is an abdication of the decision-makers' responsibility, and is a pitfall to be avoided.

On a related note, we take no pleasure in the fact that researchers feel oppressed by the need to publish in the 'top journals'. We endorse efforts to create systems that reach beyond the crudeness of the impact factor — systems that make transparent the citations and other effects of papers, and that record impacts of scientists' other work, such as their contributions to databases and the hard slog of peer review. We will develop our own systems to make such subtler credits easier to account for, and will collaborate with others working towards that same end.

*Nature's* role as a journal is complemented by its role as a science magazine. We have enhanced our journalism and externally authored opinion in recent years, and readers can anticipate further developments ahead. One example has been the introduction of a columnist. In this issue, we bid a sad farewell to David Goldston, who has been writing in these pages about the intersection of science and policy since January 2007. His Party of One column has been an essential read for US policy-makers since its debut, and earlier this year he garnered honourable mention in the US National Association of Science Writers' awards for 'The scientist delusion', his column about religion (see *Nature* 452, 17; 2008). Building on the success of Goldston's column, *Nature* will shortly introduce two new monthly columnists.

Most fundamentally, *Nature* has to reflect the values of its authors and readers. The core values of science — objectivity, independence, self-critical thinking and a relentless urge to observe, experiment and explore — are also important principles of good journalism and editing. As an unusual hybrid of magazine and journal, *Nature* can only retain readers' respect if it follows those principles while adding substantial value to the lives and work of researchers and others seriously interested in science. Our commitment to fulfil these ambitions is as strong as it has ever been. ■

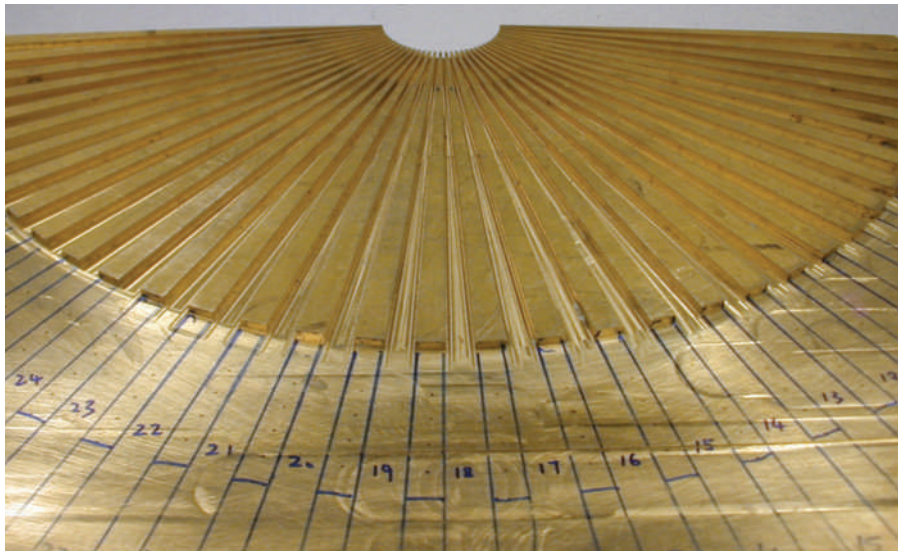
# RESEARCH HIGHLIGHTS

## Brass eye

*Nature Mater.* doi:10.1038/nmat2561 (2009)

Whether it be light or sound, there is a limit to the detail a lens can resolve. Because of diffraction, a conventional lens will never resolve features smaller than about half a wavelength.

But this rule doesn't hold for lenses made of metamaterials, composite materials that bend light or sound waves in unusual ways. Xiang Zhang and his colleagues at the University of California, Berkeley, used 36 precisely machined brass fins to create just such a lens for sound waves. Their 'hyperlens' compresses evanescent sound waves, which fade away very quickly, into more robust propagating waves, magnifying sub-diffraction features by a factor of eight. The technology could lead to higher-resolution images in sonar and ultrasound.



## ANIMAL BEHAVIOUR

### Fruit-bat fellatio

*PLoS One* 4, e7595 (2009)

Libiao Zhang of Guangdong Entomological Institute in China and his colleagues have documented what may be the first case of regular fellatio in adult animals other than humans.

They report that female short-nosed fruit bats (*Cynopterus sphinx*) licked their mate's penis during 14 of 20 observed copulations. Matings that involved licking lasted significantly longer than those that did not.

Possible functions for this behaviour include stimulation to prolong copulation and assist fertilization; mate guarding; antibacterial effects; and the detection of chemicals assisting in mate choice.

The authors say their observations could suggest a possible adaptive benefit for the activity in this species.

## CLIMATE

### Aerosols overlooked

*Science* 326, 716–718 (2009)

Aerosol particles in the atmosphere and their short-term interactions with gases have been overlooked in climate-change mitigation policies and emissions-trading schemes, say Drew Shindell from NASA's Goddard Institute for Space Studies, New York, and his colleagues.

The team found that methane's global-warming potential is greater when combined with aerosols — atmospheric particles such as dust, sea salt, sulphates and soot. The picture is complicated: some aerosols cool the

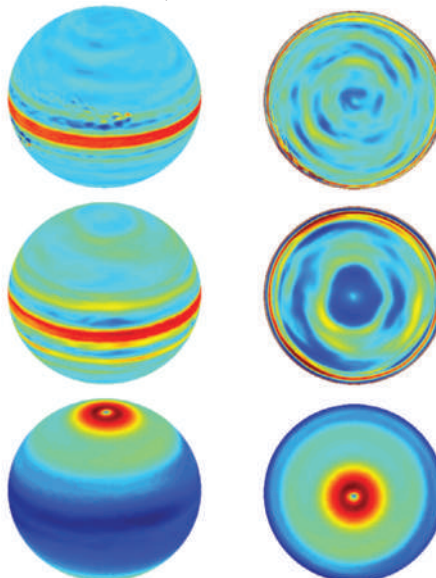
planet, some warm it, and other components of the overall climate system, such as ozone, are also affected.

## PLANETARY SCIENCE

### Jet setting

*Icarus* doi:10.1016/j.icarus.2009.10.006 (2009)

The fast-flowing gas currents that give the giant planets their banded appearance may be caused by energy released when water condenses into clouds and storms. Jupiter (pictured below, top) and Saturn (below, middle) have 20 to 30 of these jet streams, including a broad eastward jet that flows around the equator. Uranus and Neptune (below, bottom) each have only about three, with a westward jet at the equator.



Yuan Lian and Adam Showman from the University of Arizona in Tucson used simulations to show that the heat produced when water condenses during storms could drive jet streams matching those observed on all four planets. They found that expected differences in water abundance on the planets could account for the distinct stream patterns.

## NEUROSCIENCE

### Bridging the gap

*Neuron* 64, 165–172 (2009)

With the right approach, spinal cord neurons can be coaxed to regenerate more than a year after spinal injury.

The strategy, identified by Mark Tuszynski of the University of California, San Diego, and his colleagues, involves a series of steps targeting cellular growth mechanisms and environmental obstacles that hinder neurons' ability to regenerate.

The researchers severed the spinal cord in rats and applied their complex combination therapy either six weeks or 15 months later. In both cases the long axons of sensory neurons grew across the lesion and beyond, although the number of regenerating axons and the distance they travelled fell over time.

## EVOLUTIONARY GENETICS

### Mutation elevation

*Genetics* doi:10.1534/genetics.109.108803 (2009)

High mutation rates are thought to drive populations to extinction through the build-up of harmful genetic changes. This is not necessarily so, according to a new experiment.

James Bull and his colleagues at the



University of Texas at Austin exposed a type of virus called a T7 bacteriophage to a toxin that increased its mutation rate by 2–3 orders of magnitude, to about four mutations per genome each generation. After 200 generations, the virus's fitness had increased, rather than decreasing as expected. The genomes were riddled with hundreds of deleterious mutations, but 28 adaptive mutations, mostly in DNA-metabolism genes, reached high frequencies.

## MOON MATTERS

### Lunar hideaway

*Geophys. Res. Lett.* doi:10.1029/2009GL040635 (2009)

A craft orbiting the Moon has discovered an apparent 65-metre-wide opening into a deep lava tube, offering a possible site for a station from which to travel farther into space.

Junichi Haruyama of the Japanese Institute of Space and Astronautical Science in Sagami-hara and his team estimated from shadows that the tube is 80–88 metres deep, enough to shield a space outpost from meteors.

The tube is in the Marius Hills region of a volcanic province on the near side of the Moon. The SELENE polar orbiter, which operated from September 2007 until June this year, photographed the tube with a terrain camera and a multi-beam imager.

## NEUROLOGY

### Impossible movements

*Proc. Natl Acad. Sci. USA* doi:10.1073/pnas.0907151106 (2009)

Scientists can alter people's perception of their bodies by playing with their sensory input, for example by using trick mirrors or touch. Now Lorimer Moseley of the Prince of

Wales Medical Research Institute in Sydney, Australia, and Peter Brugger of University Hospital Zurich in Switzerland show that the sensation of having impossible bodily forms can be generated using thought alone.

The team asked seven amputees who have a phantom arm to perform a wrist movement with the phantom limb that would be impossible with an actual wrist. Four were able to learn the movement, which induced a change in body image and made some previously easy movements of the phantom arm more difficult.

## ANIMAL BEHAVIOUR

### Deep sleep

*Biol. Lett.* doi:10.1098/rsbl.2009.0719 (2009)

Northern elephant seals (*Mirounga angustirostris*) spend months at sea, surfacing only briefly between 20-minute dives. Yoko Mitani of Hokkaido University in Japan and her colleagues attached data recorders to six animals (see picture, below) to see if they could be resting on 'drift dives', which involve little or no active swimming.

During such dives, they found, the seals descend rapidly to at least 135 metres, then roll onto their backs and drift downwards,



wobbling like a falling leaf. This slows the descent rate significantly.

The authors suggest the seals dive quickly to below where killer whales and great white sharks normally hunt and then go belly-up to avoid sinking too far into the depths while they nap.

## ECOLOGY

### Boom and bust

*Ecol. Lett.* doi:10.1111/j.1461-0248.2009.01391.x (2009)

In an ecological system with two predator and two prey species, theory predicts two ways in which the species' population dynamics can become coupled. If both predators eat both prey, then prey populations will oscillate together, booming when predators are rare and crashing when predators are faced with a prey glut, boom in turn. But if each predator eats a separate prey, and the two prey species compete, theory says the prey populations will oscillate out of sync from one another, as first one, then the other, dominates resources.

Jef Huisman of the University of Amsterdam and his colleagues studied eight years of measurements from a Baltic Sea plankton community maintained in a laboratory. From amid the chaos of thousands of population measurements, they were able to discern for the first time in real life two coupled predator–prey cycles oscillating out of sync, showing strong effects of prey competition.

#### Correction

The Research Highlight 'Galaxy size matters' (*Nature* **461**, 1177; 2009) gave the wrong distances for the 225 galaxies surveyed by Sune Toft and his colleagues. They are actually between 4.4 billion and 5.9 billion parsecs from Earth.

## JOURNAL CLUB

**Jan Zaanen**  
Leiden University,  
the Netherlands

### A theoretical physicist journeys to a hairy black hole's horizon.

Rumour has it that Steven Spielberg is producing the ultimate science fiction movie, using state-of-the-art general-relativity simulations to create a realistic image of the warped space-time near a black hole. But wouldn't it be great to see such worlds in real life? In fact, you can: by extending your eyesight with 'AdS/CFT', a mathematical

result of string theory that describes a 'through the looking-glass' experience that would embarrass the imagination of Lewis Carroll.

AdS/CFT states that information about the strange world of the black hole is, in a very indirect way, encoded in or 'imaged' by the properties of certain quantum-weird forms of matter. Scientists realized recently that these 'quantum critical' states of matter are routinely produced in condensed-matter laboratories. But a particular prediction of AdS/CFT made the string theorists nervous: the event horizon of the special black hole that is imaged by the quantum critical

electrons seems to imply that the latter should show a macroscopic entropy at zero temperature. It has further been predicted that the black hole would be unstable and would eventually suck up 'stuff' from its surroundings, covering its horizon with 'hair' (S. A. Hartnoll *et al.* *J. High Energy Phys.* **2008**, 015; 2008). In the electron system, out of the blue and at a quite low temperature, some unexpected order will set in that removes the ground-state entropy, giving it a unique ground state.

Intriguingly, I learned the other day that condensed-matter experimentalists, unaware of the string theorists' nervousness, are

now in the grip of the same idea. The latest thermodynamic experiments on quantum-critical electrons are suggestive (albeit inconclusive) of a developing zero temperature entropy — for the experimentalists, a catastrophe — interrupted at a very low temperature by the onset of an exotic quantum liquid crystalline order (Z. Fisk *Science* **325**, 1348–1349; 2009). It may be that we don't need spacecraft or Spielberg to visit black holes, just a little patience with the condensed-matter experimentalists.

Discuss the papers at <http://blogs.nature.com/nature/journalclub>



# NEWS BRIEFING

## ● POLICY

**HIV ban lifted:** US President Barack Obama has announced the end of a 22-year-old policy barring HIV-positive individuals who are not US citizens from entering the United States. The new policy comes into effect on 4 January 2010. "The ban was not based on any scientific rationale," says Anthony Fauci, director of the National Institute of Allergy and Infectious Diseases in Bethesda, Maryland. "We are all very pleased it has been lifted. Now we will hopefully have international conferences here where people can come and go without getting into difficulty."

**University reform:** Would-be university professors in Italy will in future have to pass a national exam before being eligible to apply for positions, according to a reform bill unveiled on 28 October. Education minister Mariastella Gelmini says the aim of the legislation is to stop universities from hiring unsuitable local candidates who are owed favours. It also relegates to short-term contracts the lowest of three positions on the academic scale, known as 'researchers'. The bill may be watered down on its passage through parliament, but Gelmini hopes it will be in operation by next March.

**US-Cuba exchange:** A delegation of US scientists and environmental officials met with Cuban counterparts in Havana last week to start a new era of scientific exchange between the two countries. The talks were initiated by the Environmental Defense Fund in Washington DC, which has special permits from the US government for interactions with Cuba. Scientists from Mexico also participated. Other non-governmental organizations, the New America Foundation and the American Association for the Advancement of Science, both of Washington DC, are also organizing scientific exchange visits (see *Nature* **460**, 447; 2009).



F. LENOIR/REUTERS

## CLIMATE FUNDING CREEPS CLOSER

Warming up for climate negotiations in Barcelona this week (see our blog at [go.nature.com/TSZq6t](http://go.nature.com/TSZq6t)), European Union leaders agreed at a summit in Brussels on 29–30 October that the 27-nation bloc would pay its "fair share" of aid to help developing countries cope with and fight climate change, although the amount was not specified. The meeting estimated that, in total, roughly €100 billion (US\$148 billion) a year by 2020 would be needed, of which €22 billion to €50 billion a year would come from direct government support. "All countries, except the least developed, should contribute to international public financing," a post-meeting statement said.

## ● BUSINESS

**Changing focus:** Drug maker GlaxoSmithKline (GSK) says it has made significant progress in transforming itself from a "white pill/Western market" company into a business focused more on emerging markets and consumer health products. Fewer than 30% of GSK's third-quarter 2009 sales, reported on 28 October, came from traditional pill sales in the developed world, noted Andrew Witty, GSK's chief executive; that proportion was 38% when Witty started at the company in May 2008. GSK predicted that its H1N1 flu products would provide 'significant sales' towards the end of the year. Novartis and Sanofi-Aventis made similar predictions for their flu drugs.

**Biosimilars competition:** Legislation unveiled by the US House of Representatives last week as part of a massive US health-reform bill grants makers

of brand-name biological drugs 12 years of market exclusivity before they face competition from manufacturers of their generic equivalents, called biosimilars. A Senate version of the bill also guarantees brand-name drug makers 12 years of competition-free market access — a period that the Generic Pharmaceutical Association in Arlington, Virginia, called "unwarranted and excessive".

**Venture capital:** China's powerful National Development and Reform Commission (NDRC) has announced 20 venture-capital funds targeting high-tech enterprises. The 9-billion-renminbi (US\$1.3-billion) investment comprises 2.2 billion renminbi from central and local governments, with the remainder coming from private industry, state media said. The funds will target electronics, information technology,

## SOUND BITES

**"This should be the day on which we start to turn the tide in the battle against AIDS."**

Jacob Zuma

In a commitment welcomed by AIDS researchers, the South African president acknowledged on 29 October that his country needed a "renewed onslaught" against the disease.

Source: AP

biotechnology, drugs, new energy, conservation and environmental protection.

**R&D investment:** Swiss drug maker Novartis said it would invest US\$1 billion over five years to expand its Institute for BioMedical Research in Shanghai into the company's third-largest research and development centre (after a facility in Cambridge, Massachusetts, and headquarters in Basel, Switzerland). Announcing the plans on 3 November, the company said the centre would become the largest pharmaceutical research plant in China.

## RESEARCH

**Water satellite:** The European Space Agency successfully launched a satellite on 2 November that will map changes in Earth's soil moisture and oceanic salt levels. The Soil Moisture and Ocean Salinity (SMOS) satellite measures microwave emissions at a particular frequency and will produce global soil-moisture maps with a spatial resolution of about 50 kilometres. See [go.nature.com/DiHh2o](http://go.nature.com/DiHh2o) for more.

**Rocket test:** NASA successfully test-launched its prototype Ares I-X rocket last week, although a parachute failure meant that the discarded first-stage booster smacked into the ocean harder than planned, collecting a sizable dent. Ares I may replace the shuttle as the main means of getting astronauts to and from space — if its development is not cancelled

## NEWS MAKER



David Nutt

**The UK government sacked Nutt, the chair of its independent Advisory Council on the Misuse of Drugs, after he reiterated his views on the relative safety of various drugs. See Editorial, page 11, and [go.nature.com/WZ5c16](http://go.nature.com/WZ5c16) for more.**

following the report of an independent presidential panel (*Nature* 461, 153; 2009).

**Antarctica drilling:** Russia has postponed for another year its plans to drill into Antarctica's sub-glacial Lake Vostok (*Nature*, 454, 258–259; 2008). Entry into the uniquely pristine lake, 3,750 metres below the Antarctic ice

sheet, is now planned for the 2010–11 drilling season, said Valery Lukin, director of the Russian Antarctic Expedition. See [go.nature.com/EkyjYM](http://go.nature.com/EkyjYM) for more.

**Rocket scientist dies:** Qian Xuesen, widely known as the father of China's nuclear-missile and space programmes, died on 31 October at the age of 98. A prominent scientist involved in setting up the Jet Propulsion Laboratory in Pasadena, California, Qian was detained in the United States for five years on suspicion of being a communist spy, before being deported in 1955. In China, he established the country's first ballistic-missile programme. His research led to nuclear-weapons testing in the 1960s and the launch of China's first manned space mission in 2003.

## FUNDING

**Stem-cell grants:** The California Institute for Regenerative Medicine in San Francisco has awarded US\$230 million to 14 research teams, with the intention of helping to move stem-cell therapies into the clinic within four years. The grants, which coordinate basic and clinical researchers, target a range of diseases, including brain cancer, stroke, heart attack, macular degeneration and sickle-cell anaemia. See page 22 for more.

**Smart grid:** The US Department of Energy bolstered efforts to establish a 'smart' electric grid by announcing on 27 October US\$3.4 billion in grants for various grid modernization

## THE WEEK AHEAD

### 6–15 NOVEMBER

**The International Commission for the Conservation of Atlantic Tuna meets in Recife, Brazil. With the tuna population in dramatic decline, the commission is under pressure to cut fishing quotas and clamp down on trade.**

► [www.iccat.int/en](http://www.iccat.int/en)

### 9 NOVEMBER

**Marking the 20th anniversary of the fall of the Berlin Wall, researchers will gather in the German capital to discuss future breakthroughs in fields ranging from fusion to infant development, at the 'Falling Walls' conference.**

► [www.falling-walls.com](http://www.falling-walls.com)

### 10–11 NOVEMBER

**Accra, Ghana, hosts the fifth annual international conference of the African Science Academy Development Initiative.**

► [go.nature.com/FCYMPZ](http://go.nature.com/FCYMPZ)

### 10 NOVEMBER

**The International Energy Agency launches the 2009 edition of its annual World Energy Outlook in London.**

► [www.worldenergyoutlook.org](http://www.worldenergyoutlook.org)

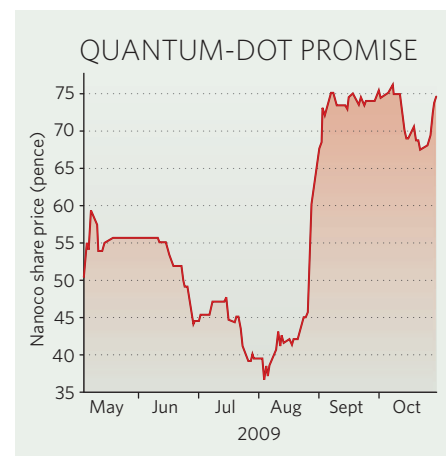
## BUSINESS WATCH

Nanoco, a UK spin-off company that makes quantum dots (semiconductor nanocrystals used in displays and bioimaging) has released its first full-year financial results since floating on the London Stock Exchange's Alternative Investment Market (AIM) in May. At £1.99 million (US\$3.26 million), Nanoco's revenues in the year to 31 July 2009 were up from 2008 levels of £1.08 million, although the company still made a loss of £780,000, compared with £698,000 last year.

In September, Nanoco signed an agreement with an unnamed Japanese electronics manufacturer to make quantum dots for use in light-emitting diodes for television backlighting.

This deal and one struck in December 2007 with another Japanese display company are more significant than the AIM listing, says David Hwang, from market-analysis company Lux Research in New York. "It pushes them out of the little leagues," he says, adding that partnerships with major electronics corporations are crucial.

The quantum-dot market is still small, but ambitious (*Nature* 459, 760–761; 2009). The industry's bigger players include private companies QD Vision of Watertown, Massachusetts, and Nanosys in Palo Alto, California; Life Technologies also has a quantum-dot division, primarily for bioimaging.



SOURCE: AIM

## NEWS

# Brazil mulls major climate action

If adopted, the move would put the country ahead of other developing nations on emissions curbs.



Building on an existing pledge to slash deforestation rates in the Amazon, Brazil is considering a commitment to substantially reduce cumulative greenhouse-gas emissions over the next decade.

Details were under discussion as *Nature* went to press, but such a commitment would represent the most significant step yet by a developing country going into December's United Nations climate summit in Copenhagen. It could also help to bridge the gap between rich and poor countries, while increasing pressure on industrialized nations to boost their commitments.

"What's really important is that Brazil has seen that it can lead by example," says Steve Schwartzman, who heads tropical-forest policy for the Environmental Defense Fund in Washington DC. "This is a major step. It's tantamount to taking a national cap."

Under the current global-warming framework, only industrialized nations would be required to 'cap' and then reduce their greenhouse-gas emissions. Developing countries are focusing on policies to reduce emissions growth and set up low-carbon development.

Although Brazil is not discussing a binding emissions cap, scientists and government officials have been updating the nation's greenhouse-gas inventories and quantifying potential reductions. Preliminary results from one analysis suggest that the nation could cut its projected emissions in 2020 by 35%, which would equate to roughly 8% below 2007 levels. The government is also discussing even deeper cuts, all contingent on international aid.

"We are just estimating the magnitude of reductions that we can obtain by 2020," says Carlos Nobre, a climate scientist at Brazil's National Institute for Space Research in São José dos Campos. "The next step is to put a dollar figure to each of those things, and then one has to discuss how to get financing."

Working with the scientist network Rede-Clima, Nobre recently calculated for Brazil's environment ministry that the country could reduce its annual emissions by the equivalent of nearly 1 billion tonnes of carbon dioxide (see graphic) by 2020. Nearly two-thirds of that would come from an existing commitment to an 80% reduction in deforestation — now

responsible for more than half of Brazil's total emissions — by 2020. But Rede-Clima estimates that another 320 million tonnes could come from the energy and agriculture sectors, as well as reforestation efforts.

"If we get assistance and we develop good policies here, we might get to 2020 with a maximum 40% reduction," Nobre says, and that would cost several billion dollars or more.

Marina Silva, an environmental advocate who resigned as Brazil's environment minister last year amid internal opposition to her policies, has been using her new post in the Senate to push for a comprehensive national target in Copenhagen. Speaking through an interpreter in Washington DC last week, she said

the principle of "common but differentiated responsibilities" might require wealthy nations to take the lead but does not excuse developing countries from taking action.

"It cannot be something that the developing countries hide behind in order to avoid making reductions," says Silva, who is expected to run as the Green Party candidate for president next year. "Our planet does not take into account where the emissions are coming from."

The government is also rethinking its position on the role of forest carbon in a future climate treaty. To date, the country has taken a lone stance in demanding that wealthy countries should not be able to offset their greenhouse-gas emissions by funding forest conservation in the tropics. The fear is that offsets would allow industrialized countries to delay action at home, but observers say that position may soften.

Brazil created the Amazon Fund last year as an alternative mechanism that would allow donor countries to help pay for the country's ongoing programme to curb deforestation (see *Nature* 460, 936-937; 2009). Norway has pledged US\$1 billion up to 2015, assuming Brazil continues to make progress.

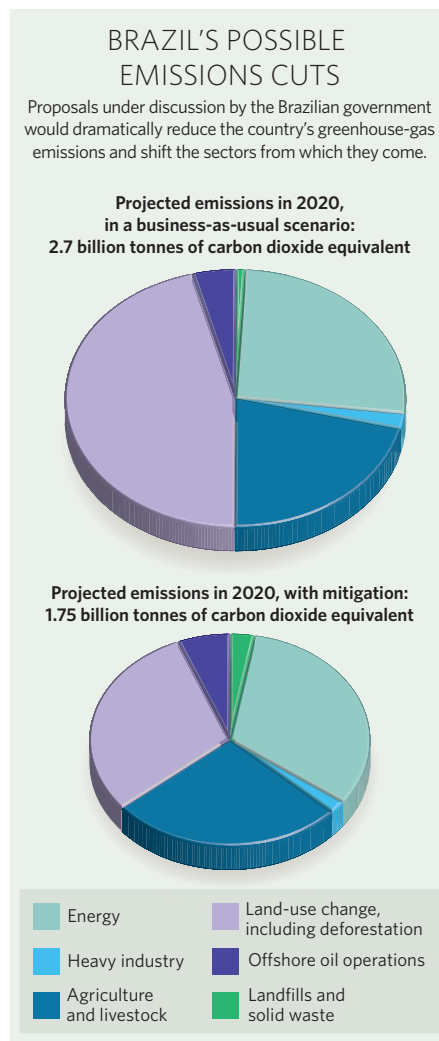
But President Luiz Inácio Lula da Silva appointed a task force to study the issue after nine governors representing Amazon states urged him in June to reconsider Brazil's policy on carbon markets, which they said represent "a golden opportunity" for sustainable development and forest conservation. That panel has proposed allowing nations to offset up to 10% of their commitment by purchasing carbon credits for avoided deforestation.

Schwartzman credits Silva with initiating change within the federal government. "Not very long ago, the official Brazilian position was that they did not want any discussion of deforestation in the international negotiations," he says. "It was at her urging that the Brazilian government reopened that policy and came to a different conclusion."

Officials with the environment ministry could not be reached for comment. As *Nature* went to press, Lula had scheduled a 3 November cabinet meeting to discuss the issue. It was unclear whether a formal position would be announced this week in Barcelona, where climate negotiators are meeting for the last round of formal talks before Copenhagen. ■

Jeff Tollefson

See also [www.nature.com/roadtocopenhagen](http://www.nature.com/roadtocopenhagen)



SOURCE: REDE-CLIMA





**AEROSOLS MAKE METHANE MORE POTENT**  
Air pollution linked more closely to climate concerns.  
[go.nature.com/ARFpmf](http://go.nature.com/ARFpmf)

R. KRAUSE/REUTERS

# Initiative targets malaria eradication



Next-generation vaccines are intended to move beyond the initial infection phase.

The PATH Malaria Vaccine Initiative (MVI), the main public-private partnership for developing vaccines against the disease, this week announced a major overhaul of the sorts of vaccine candidates it will support in the future, as well as the ways in which it will select them.

The move follows a 2007 call by Bill and Melinda Gates — major funders of malaria research — for researchers to not just control the disease, but aim to eradicate the parasite. “Whether you believe that eradication is attainable or not, it has prompted substantial discussion in the malaria community as to what tools would be needed to reach it,” says Robert Sinden, a malaria biologist at Imperial College London.

Until now, clinical efforts towards a vaccine have focused on the individual level by reducing cases and deaths. But the MVI, based in Bethesda, Maryland, will now also invest heavily in vaccines that block transmission of the malaria parasite. These vaccines provide no direct benefit to individuals but might fight the disease at the population level.

Transmission-blocking vaccines generate antibodies to the stage of the parasite that replicates sexually in the mosquito gut. So when a mosquito takes blood from an infected person who has been vaccinated, it will also suck up antibodies and immune cells that destroy the parasites as they replicate. Laboratory tests and pilot studies in humans have shown that such vaccinations drastically lower the numbers of infected mosquitoes.

“I’m delighted that the MVI is finally

plugging this gap in its portfolio,” says Adrian Hill, a malaria-vaccine researcher at the University of Oxford, UK.

Another type of vaccine targeted for increased support is live-attenuated vaccines of the *Plasmodium* sporozoite — the stage of the parasite that initially infects the liver. For decades, the logistics of making a working vaccine using sporozoites have been considered insurmountable; among other things it requires manually dissecting billions of sporozoites in mosquito farms.

Sanaria, a company in Rockville, Maryland, run by veteran malaria researcher Stephen Hoffman, took up the sporozoite challenge six years ago, and in April it began a phase I clinical trial of its radiation-attenuated vaccine. “A lot of people were sceptical about that, and I think a lot of credit should go to Hoffman and his people at Sanaria [because] they have overcome that,” says Anthony Fauci, director of the US National Institute of Allergy and Infectious Diseases. The trial will end next June.

Even if the trial is a success, however, the obstacles to scaling-up and delivering the vaccines remain formidable. Nonetheless, Sanaria’s progress has spurred the MVI to take the approach seriously. It also intends to support several labs that are developing sporozoites attenuated by knocking out genes. Although still at the research stage, this technique might be a more reliable way to attenuate the parasite than radiation, says Ashley Birkett, the

MVI’s director of preclinical development.

By 2025, the MVI aims to have next-generation vaccines that are more than 80% effective and that last for at least four years. The most effective first-generation effort in its portfolio, GlaxoSmithKline Biologicals’ RTS,S vaccine — which targets the initial infection phase in humans — falls far short of that. Phase II trials have suggested that it cuts infections and severe disease by around 50%, enough to save many lives, although the duration of the protective effect remains uncertain. A phase III trial is under way at 11 sites in Africa.

Combining RTS,S with vaccines targeted at other stages of the malaria parasite’s life cycle could produce much stronger effects. Animal studies suggest that combined vaccines show a substantial multiplier effect, says Hill, whose T-cell vaccines are a current favourite for the first trials with RTS,S. The MVI intends to pursue only combined vaccine trials in future.

Moreover, the RTS,S results have set a higher bar for contenders, says Birkett. In the future, he says, vaccine candidates will need to pass much more extensive preclinical testing, and initial testing in a few human volunteers, with only the highly effective ones being taken forward. Such early testing systems have been chronically underfunded.

One approach that is proving powerful is human challenge experiments, in which a few volunteers are vaccinated and their immune responses closely studied. Such studies should be the rule, says Hill. “Nothing should be taken forward into Africa without knowing it does something.”

Birkett says that the MVI is supporting a dedicated Human Challenge Center at the Seattle Biomedical Research Institute in Washington, one of only four such centres worldwide, that will open next year.

As well as weeding out potential vaccine candidates, the MVI intends to look at funding partnerships beyond the malaria-vaccine community to find feedstock for its pipeline. “Of course the malaria community will remain our core community,” says Birkett, “but there’s a lot of untapped expertise out there, including experts in HIV, influenza, cancer and biotech who have different vaccine strategies.”

Declan Butler

D. POLAND/PATH MALARIA VACCINE INITIATIVE

# Children's study fights to survive

US politicians, once supportive of a massive research project on childhood health, are now criticizing it.

## NORTH CAROLINA

Late last month, in a former Blockbuster Video store in tiny Kenansville, North Carolina, an insulated cardboard box awaited a placenta. Collected after a nearby birth the night before, the placenta would be shipped to the University of Rochester Medical Center, home of the central histological lab for the US National Children's Study. There, pathologists would catalogue the placenta's weight, size and shape; note any abnormal features; fix it in formalin and embed it in paraffin; then make slides and tissue blocks to preserve it for future use by researchers studying major childhood ailments from asthma to obesity.

Many more specimens may soon be coming from rural Duplin County, North Carolina. Over the next several years, investigators in this land of hog farms and cotton fields hope to enrol half the new pregnancies in what is arguably the most ambitious study ever of how the environment affects children's health.

Authorized by Congress in 2000, the study has ramped up only recently, with funding rising from US\$69 million in 2007 to \$180 million this year (see graph). It aims to study more than 100,000 children at 105 locations around the country, following them from before birth to age 21. Its 28 proposed hypotheses intend to explore, among other things, whether prenatal infection is a risk factor for autism; how media exposure influences neurological, cognitive and social development; and how recurrent, low-level pesticide exposure might affect cognitive skills.

The study is now enrolling patients at seven 'vanguard' centres, including the one in Duplin County. But just as it is getting off the ground, it is running into political tumult. In August, angry senators on the Senate appropriations committee declared that the National Institutes of Health (NIH), the study's principal overseer, had committed a "serious breach of trust" by withholding from Congress news that estimates of the study's cost — initially around \$3 billion over 25 years — had grown to as much as \$7 billion. Among other things, officials at the NIH's National Institute of Child



A study of 100,000 US children should provide insights on how the environment affects their health.

Health and Human Development (NICHD) had failed to include estimates of the indirect costs routinely paid to institutions.

Until the NIH provides "the most up to date information possible", the senators wrote, they would reserve a decision on how many dollars — "if any" — they would allocate to the study in 2010. (The number will be settled during upcoming negotiations between appropriators in the Senate and the House of Representatives; lawmakers in the latter have agreed to fund the study at \$194 million in 2010.)

In the staid parlance of Capitol Hill reports, the words "breach of trust" landed with the impact of a hand grenade. "I have never seen report language referring to anything that the

NIH has ever done as a 'serious breach of trust'. I think that is telling," says a seasoned observer of biomedical politics in Washington DC who asked not to be named.

The NIH had already responded by shaking up its staff, removing study director Peter Scheidt in July. By early October, NICHD director Duane Alexander was out of his job, too. (Alexander, now at the NIH's Fogarty International Center, has said that he was already considering leaving the director's job after 23 years.)

The NICHD's new acting director, Susan Shurin, says that "we obviously take the Senate's concerns very seriously." Some of the 28 hypotheses will "probably" be dropped, she says; "we are going to be managing to a budget." The NIH's deputy director Raynard Kington says that when a pilot phase finishes in spring 2010,

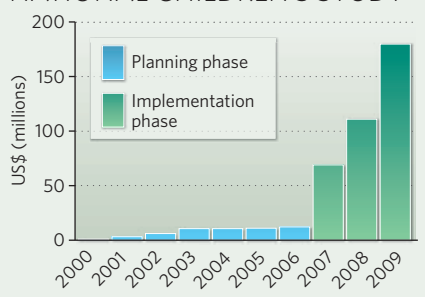
the agency will embark on a "rigorous evaluation" to determine the study's future trajectory.

The Senate rebuke was far from the study's first labour pain. In the mid-2000s, the White House under George W. Bush had repeatedly zeroed out its budget — with the support of the NIH's then-director Elias Zerhouni, who argued that its long-term costs were prohibitive. But members of Congress, many of whom had potential study locations in their districts, repeatedly reinserted the funding.

## Power of the people

Study leaders say that big is the way to go with this study — that the power of 100,000 participants and the extensive amount of information they will provide, from house-dust samples to parental toenail clippings, will produce a data bank that detectives of disease causation can mine for decades to come. Other major longitudinal studies, such as the Framingham Heart Study, have provided this kind of trove — but for adults, notes Barbara Entwisle, director of the Carolina Population Center in Chapel Hill and the principal investigator at the Duplin County centre. "All of the recent science shows the importance of what happens early on in people's lives for health all the way through their lives," she says. "This is the future of our country."

FUNDING FOR THE NATIONAL CHILDREN'S STUDY







# 'BAD GENES' WIN LIGHTER SENTENCE FOR MURDERER

Jail term cut after genetic tests.

[go.nature.com/zpUfSg](http://go.nature.com/zpUfSg)

A. T. SUMNER/GETTY IMAGES

In 2004, after long and controversial consultations, study leaders adopted a household-based sampling strategy, a costly and labour-intensive decision meaning that participants from the 105 far-flung study locations will be more representative of the broad population than, say, those recruited from health facilities and doctor's offices, a more common approach.

In May 2008, in a 140-page report, the National Research Council and the Institute of Medicine praised the study's household-based sampling approach, as well as the statistical power rendered by its 100,000 participants. But it also identified a laundry list of weaknesses in the study's design. The first was the absence of a pilot phase; because the vanguard centres' data were to be part of the final data set, investigators would be less likely to experiment with methods and change things that weren't working. That concern is being addressed, says Shurin, by turning the vanguard centres into long-term development platforms for the main study's methodology and feasibility, and not, as initially intended, folding them into the main study. Study leaders have also pushed back the launch of the study's main phase until late 2011, giving investigators time to tweak the methodology according to what's learned at the vanguard centres.

Despite the rough waters, advocates remain passionate about the study's potential, arguing that it is an investment that will more than pay for itself. Leonardo Trasande, a study investigator at the Mount Sinai School of Medicine in New York, notes that just six of the conditions the study explores — asthma, autism, diabetes, injury, obesity and schizophrenia — cost the United States at least \$650 billion a year (L. Trasande and P. J. Landrigan *Environ. Health Persp.* **112**, A789–A790; 2004). "If the study identifies preventive interventions that can reduce those annual costs by even 1%, [a \$3-billion] study repays itself twofold in one year," he says.

Meanwhile, Duplin County continues to enrol participants. While recruiters work church dinners and organize soccer tournaments, data collectors recently finished enumerating 10,800 households — roughly half of those in the county — aiming to locate reproductive-age women.

"I have two daughters. Both have asthma. I have lived this," says Entwisle. "But the scientific reason for my involvement has to do with the fact that there are so many questions that need answers. This design is exactly what needs to be done to begin to develop some of those answers."

**Meredith Wadman**

## 10,000 genomes to come

Results are just beginning to arrive for the 1000 Genomes Project, a genomic study of human diversity. However, an international group is already planning something even more ambitious — a 10,000 genomes project. The initiative, called Genome 10K, aims to tackle thousands of vertebrate species.

Project members write this week in the *Journal of Heredity* that the effort will provide an unprecedented look at the genomic mechanisms for generating diversity in a group of animals with different lifestyles and adaptations.

"We see this stunning diversity of forms in vertebrates, from manatees to anteaters,"



The anteater could have its genome sequenced.

says David Haussler of the University of California, Santa Cruz, one of the three masterminds behind the project, which involves 68 scientists from five continents. "What continually strikes me is the unbelievable malleability and adaptability of the vertebrate genome," he says, "and we have an enormous amount to learn about the genetic roots of that."

Haussler, Stephen O'Brien of the National Cancer Institute in Frederick, Maryland, and Oliver Ryder, director of genetics at the San Diego Zoo's Institute for Conservation Research in Escondido, California, decided to organize the project after realizing that one of the major obstacles to non-human sequencing projects has been collecting and organizing specimens. After an April meeting in Santa Cruz to bring in other scientists, the team now has a database of samples from more than 16,000 species from 50 institutions. The scientists are also planning a pilot project to sample portions of the genomes of a small subset of these species.

The group is looking for funding for the main phase of the project, which could cost anywhere from US\$10 million to \$100 million, depending on the costs to process and sequence each sample. The team anticipates that sequencing costs will drop below \$10,000 per genome within a few years, making it feasible to sequence the entire genomes of 10,000 vertebrates within this budget.

However, analysing all the data will be an enormous challenge, because it is still not easy to assemble new genomes from the short readouts of DNA delivered by current sequencing technologies.

Other scientists say that the project

sounds exciting, but question the decision to sequence so deeply solely among vertebrates.

David Maddison, who studies beetle phylogeny at Oregon State University in Corvallis, points out that, so far, only one beetle species has had its genome sequenced, despite there being about six times as many beetle species as there are vertebrates.

"My biggest concern is that if one were to decide where our funds

should go — to 10,000 vertebrates or more generally scattered across organismal diversity — I would very strongly argue for the latter," says Maddison, who conceived the Tree of Life project, a web-based catalogue of biodiversity.

O'Brien agrees that invertebrate-genome sequencing is "a valuable area that should also be considered for whole-genome-sequence assessment". Other scientists are already planning a large invertebrate sequencing project.

Indeed, O'Brien expects Genome 10K to be the first study in what he anticipates will be a larger shift toward using sequencing technology to study biodiversity. He compares it to the culture shift that accompanied the invention of the printing press, which was first used to print the Bible and then for broader purposes.

"If the Human Genome Project is the Bible," he says, "then the [Genome 10K] is a library that gets filled up with other books."

Erika Check Hayden

A. SHAH/NATUREPL.COM



# California stem-cell grants awarded

First major round of research targeted at therapies takes off.

The starting gun has fired for 14 research teams, based in California, who now have four years to make good on the therapeutic promise of stem cells.

On 28 October, the California Institute for Regenerative Medicine (CIRM) committed US\$230 million to teams of basic and clinical researchers aiming to move experimental stem-cell treatments into an investigational new-drug filing with the US Food and Drug Administration. Britain and Canada together paid an additional \$43 million for four of the grants, which will include work by researchers in those countries.

Robert Klein, chair of CIRM's board, calls the awarding of the grants a "historic day".

It is the most highly anticipated round of research grants yet awarded by the San Francisco-based agency, which was set up by a 2004 ballot measure to fund stem-cell research. At the time, federal funding for research on human embryonic stem-cell lines was restricted to the relatively small number of cell lines in existence in August 2001.

Scientists outside the state say that the grants could help the stem-cell field as a whole. Sean Morrison, director of the University of Michigan Center for Stem Cell Biology in Ann Arbor, calls the awards "exciting".

"Developing an effective new treatment is a hard problem, and most good ideas won't lead to cures," he says. "But, without the resources from these disease-team grants, it would take much longer to distinguish between the good ideas that really work and the good ideas that don't quite work."

Other researchers have welcomed the awards, but note that many of the projects test ideas that are similar to work being funded elsewhere. The National Institutes of Health, for instance, is gearing up to fund research on new lines of human embryonic stem cells. It plans to spend nearly \$1 billion this year, not including money from the economic stimulus package, on research on adult and embryonic stem-cell lines in humans and other animals.

"The general [new-grant] portfolio strikes me as being similar to what is going on elsewhere,"

says haematologist Stuart Orkin of the Children's Hospital Boston in Massachusetts. "I don't see anything radically different from what I see people thinking about in other institutions, but it's great to have the funding to do it."

For instance, two of the grants will fund work to develop monoclonal antibodies — targeted biological drugs that are already approved

of California, Los Angeles, who developed the breast cancer drug trastuzumab, sold by Genentech of South San Francisco — aim to target and destroy cancer stem cells, whose existence is still a matter of some debate.

Another grant will attempt to use induced pluripotent stem cells — which have many of the capabilities of embryonic stem cells but can be tailored to match individual patients — to treat a rare skin disease called epidermolysis bullosa.

Four other awards focus on modifying adult stem cells to treat HIV and brain tumours.

Observers say that the chances of one or more of the experimental therapies making it to the clinic are improved because many of them take similar approaches to techniques that have already been approved, and because many focus on adult stem cells.

Major scientific questions remain about using embryonic stem cells in humans. The first company intending to begin clinical trials involving human embryonic stem cells — Geron of Thousand Oaks in California — has faced repeated

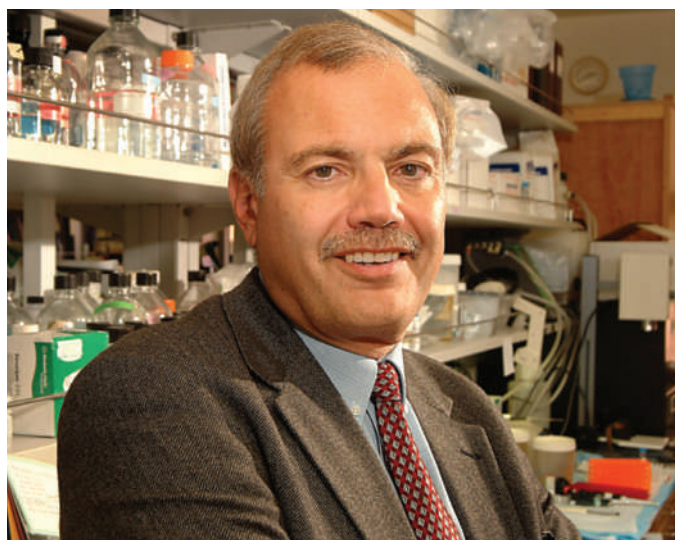
setbacks in getting its trial going.

"For adult stem cells, moving from the lab to a clinical trial doesn't take nearly as long, and four years is not outrageous at all," says Robin Young, a Philadelphia-based publisher of an orthopaedics newsletter and organizer of an annual stem-cell meeting in New York. "I would say it's very likely that some of those are going to be in clinical trials by then."

Even therapies that aim to use adult stem cells are no sure thing. In September, Osiris Therapeutics of Columbia, Maryland, announced that its treatment made from cultured adult stem cells to treat graft versus host disease had failed.

However, a success by one of the grantees could be enough for CIRM to gain voters' approval to extend its 10-year lifespan beyond its initial approved expiry date of 2014 — or 2016, if counting from when CIRM funded its first grants. "Even if one of these ideas succeeds, it would be a huge success for CIRM that would justify all of their investment in this area," Morrison says.

**Erika Check Hayden**



Dennis Slamon leads one of the teams to receive funding from CIRM.

**"For adult stem cells, moving from the lab to a clinical trial, four years is not outrageous at all."**

for many indications — to target cancer cells. Another grant will try to use a patient's own cardiac stem cells to repair damage from heart attacks, a controversial approach that is already being tested in patients. A fourth grant aims to modify patients' bone-marrow cells to correct the genetic defect that causes sickle-cell anaemia, then implant the cells back into patients.

A similar approach has been used to treat severe combined immunodeficiency disorder. "That would have been called gene therapy before, instead of stem-cell therapy, and there are a number of people doing that," Orkin points out.

## Broad disease range

Four of the grants involve work on human embryonic stem cells, intending to treat diabetes, macular degeneration, amyotrophic lateral sclerosis and stroke. And three of the cancer grants — including one co-led by Dennis Slamon, an oncologist at the University

UCLA

# Dark-matter test faces obstacles

A group of scientists is hoping to replicate a controversial Italian experiment that claims to have detected dark matter. But they might have to do so without the help, or the equipment, of the original group.

Dark matter is thought to make up around 85% of the matter in the Universe, but it rarely interacts with regular matter except through the force of gravity. Researchers working on the DAMA experiment at the Gran Sasso National Laboratory near L'Aquila, Italy, claim they have spotted direct signs of it.

The detector used by the DAMA team consists of 250 kilograms of ultrapure sodium iodide crystals placed 1,400 metres beneath Gran Sasso mountain. Over the past decade, the researchers have collected data showing that nuclei in the crystals

periodically release flashes of light, which could be caused by interactions with dark matter. Crucially, the number of flashes varies with the seasons, which would be consistent with Earth's motion through a galactic dark-matter stream (R. Bernabei *et al. Eur. Phys. J. C* 56, 333–355; 2008).

But other detectors have so far failed to see an effect, leading some to conclude that DAMA's signal is the result of radioactive contamination inside the sodium iodide crystals. "There are very good reasons to disbelieve the signal," says Adam Falkowski, a theoretical physicist at Rutgers University in Piscataway, New Jersey.

Still, Frank Calaprice at Princeton University in New Jersey says that the signal is significant enough to be followed up. "It could be right — they're careful people," he says of the DAMA team. "I

think it deserves to be checked."

Calaprice and some of his colleagues believe that building a similar detector out of sodium iodide is the only way to verify DAMA's claim definitively. They want to place those crystals inside another Gran Sasso experiment, called Borexino, which uses 300 tonnes of a liquid organic compound to detect solar neutrinos — nearly massless particles that stream from the Sun. Detecting neutrinos requires careful control of background signals, which makes Borexino the perfect place for checking whether DAMA's crystals really are seeing

dark matter, says Calaprice, a Borexino team member.

DAMA's principal investigator disagrees. The proposal "is not good for many scientific reasons", says Rita Bernabei, a physicist at the National Institute of

Nuclear Physics in Rome. "Borexino cannot ensure the fulfilment of all the stringent requirements necessary for a reliable measurement in this field," she says.

Bernabei adds that no one at Borexino has formally asked DAMA for help in setting up another sodium iodide crystal detector, and declined to say whether such help would be provided if asked.

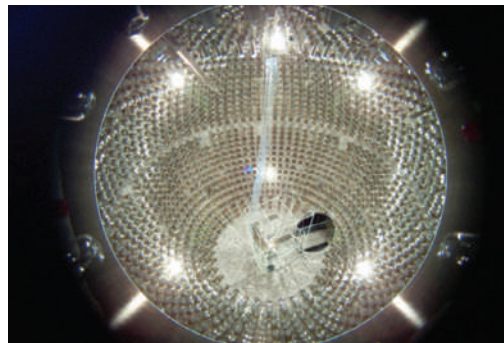
Reproducing DAMA's experiment without its cooperation is difficult because an intellectual-property agreement between the group and its supplier, Saint-Gobain in Paris, prevents the sale of the ultrapure sodium iodide crystals to other groups. A Spanish experiment known as ANAIS, located at the Canfranc underground laboratory near Zaragoza, Spain, has already been prevented from buying the crystals. "We have had to spend time

and extra money to search for an alternative," says José-Angel Villar of the University of Zaragoza, who is leading the collaboration. With the help of an American supplier, he says the group hopes to begin checking DAMA's findings soon.

Calaprice says he is hopeful that his team's test can go ahead, and is investigating the possibility of growing his own crystals. The group's proposal will have to be submitted to the full Borexino collaboration, which will decide whether to take it forward. ■

Geoff Brumfiel

**"Borexino cannot ensure the fulfilment of all the stringent requirements necessary."**



The Borexino experiment may start looking for dark matter as well as solar neutrinos.

# Science favoured by German coalition

German scientists have little to worry about in the recent coalition agreement of their government, which is planning few changes in research strategy and a lot more money. Budgets in education and research are set to almost double over the four-year legislative period.

The major coalition partner is the centre-right alliance between the Christian Democratic Union and the Christian Social Party; the junior partner is the business-friendly Free Democratic Party. Annette Schavan (Christian Democrats) remains in charge of the ministry of research and education, and Thomas Rachel is likely to remain her deputy for research.

The generous funding agreements for universities and research agreed by the previous government last June will be maintained. So the budgets of major research organizations, including the Max Planck Society and the DFG, Germany's main research granting agency, will rise by 5% a year until 2015. Universities will receive a €5-billion (US\$7.4-billion) supplement

between 2011 and 2015 to improve teaching and research.

The Excellence Initiative, a competition between universities to win 'elite' status, keeps the €2.7 billion promised for the period 2012–17. The four-year, €14.6-billion 'high-tech' programme that promoted applications in areas such as energy, climate, health and security — and that would have expired this year — will be extended and expanded.

The agreement also promises special attention to materials research, and new legislation to give research institutions more control over their budgets.

There is less clarity in two sensitive research areas. The agreement stresses that green biotechnology is valuable for the national economy, but dodges specific mention of field trials of genetically modified crops which, although legal, have become nearly impossible to carry out owing to public resistance.

The agreement similarly acknowledges the

'opportunities for health' of stem-cell research, but does not tackle the thorny issue of the restrictive German laws on human embryonic stem-cell research. Instead, it says it will consider establishing a national 'dialogue platform' for more public discussions. The Free Democrats have withdrawn from their pre-election lobbying to further relax rules limiting the cell lines that German researchers may use.

"Relaxing rules to match those of some other big players in Europe, such as Switzerland, would make things a lot easier for us because we collaborate so much with European partners," says Oliver Brüstle, a stem-cell researcher at the University of Bonn.

"We are mostly satisfied, particularly by the increases in budget," says Jörg Hacker, head of the Berlin-based Robert Koch Institute and president-elect of the Leopoldina, Germany's national academy. "We hope that it will all work out as the coalition agreement plans." ■

**Alison Abbott**



# In which we say goodbye

Our departing columnist **David Goldston** reflects on some misconceptions about science and politics.

**T**his column is a bit different from the 33 that have preceded it because it is my last in this space. I recently became the director of government affairs at the Natural Resources Defense Council (NRDC), an environmental advocacy group. To avoid any potential conflicts of interest for either *Nature* or the NRDC, I, sadly (at least from my viewpoint), have to give up this column. But at least a final column affords an opportunity to sum up some thoughts on politics and science in the United States.

This column has returned repeatedly to four themes — themes that may be at odds with the way many scientists (and others) think about the political arena. The first is that politics is a legitimate part of decision-making on scientific issues. The term ‘politics’ is almost always invoked, even by politicians, as a way to portray an action or opinion as contemptible — as a reason that a policy-maker isn’t ‘just doing what is right’. But the only way a democracy can make choices or set priorities is through politics; there is hardly ever one ‘right’ answer — one that makes sense regardless of one’s goals, values, interests and ideology. (There may, however, be wrong answers — solutions that just won’t work as advertised.)

Elected officials are supposed to weigh up all those factors in making decisions, even about science. For example, although I think that it would be foolish in the extreme for Congress to eliminate social-science funding at the National Science Foundation or to cut off funding at the National Institutes of Health for research on sexual behaviour — to take two recent proposals — it is not inherently illegitimate for such proposals to be offered, and they shouldn’t be dismissed as mere instances of politics ‘interfering’ with science. Public officials need to make choices about what areas of research merit taxpayer funding; deciding what to fund is not the same as dictating what research should conclude. Moreover, opponents of particular research topics are not usually insincere or just pandering to voters. They need to be countered not with smugness or charges of politicization, but with genuine arguments about how the research they oppose can benefit the nation.

The second theme is that the public generally holds science and scientists in very high regard, as both polling and behaviour repeatedly



## PARTY OF ONE

show. This does not mean, of course, that the public understands science or always accepts scientific conclusions, evolution being an obvious case in point. But it does mean that scientists should not be quick to dismiss, say, ethical concerns about stem-cell research or queasiness about the implications of genetic or neuroscience findings as signs that the public is benighted or ‘anti-science’. Indeed, the US public often gives too much credence to anything that is labelled as ‘scientific’. That’s the reason every tentative new scientific finding, especially if health related, gets touted; it’s why tobacco companies used to have doctors appear in their cigarette ads. Credibility is scientists’ to lose.

The public esteem for science also leads politicians to describe their positions as ‘following the science’ even when the science is ambiguous or irrelevant. In a polarized political climate, arguing that your position is the only scientific one is perhaps the only way to appear to be above the political fray. This creates a perverse incentive to conflate scientific questions and policy questions, to the detriment of both: charges of ‘junk science’ get thrown around and policy questions don’t get fully debated.

So the third theme has been that science can never be the sole determinant of a policy decision, and both politicians and scientists need to distinguish between scientific disagreements and policy disagreements when debating issues. For example, questions about how much risk a substance poses need to be differentiated from questions about how much risk the public should accept.

This is easier said than done. The line is

not always clear, and the inclination to blur it is often unconscious. That blurring reflex needs to be checked, in part, by imposing requirements that would force policy-makers to make clear what they are actually arguing about. An August report from the Bipartisan Policy Center called *Improving the Use of Science in Regulatory Policy* ([www.bipartisan-policy.org/projects/science-policy](http://www.bipartisan-policy.org/projects/science-policy)) proposes, among other things, requiring the notices that announce a new regulation to list the science and policy questions that needed to be answered in designing that regulation. The report was written by a group that included both officials who had served in the administration of President George W. Bush and their antagonists. (I was the project director for the report.) Charges that ‘the science’ has been ignored in setting policy need to be scrutinized; science is not always what’s at issue.

If scientists too often think that science is being ignored in some areas of policy-making, they are too quick to think that concerns about science are driving the policy discussion in others. When it comes to debates about science policy — about the conduct or financing of research — the tendency is to interpret policies as barometers of public support for science when they are not. The fourth theme is that not everything that happens politically to science happens because of what politicians think about science.

The size of the federal science budget, for example, says far more in any given year about the overall level of federal spending than it does about attitudes towards science. In the last years of the administration of George W. Bush, research funding stagnated because of battles over total domestic spending even though both the White House and Congress supported greater science funding. The recent jump in funding certainly reflects the pro-science attitudes of President Barack Obama, but the money never would have been appropriated if the recession had not made massive government spending look more politically acceptable.

What these themes add up to, I guess, is a plea to scientists both to engage in the policy making process and to approach the political process (if not all its participants) with respect. That process is open to scientific guidance, and over time science shapes policy. But it is always valuable to keep in mind a question that every civil servant and congressional staffer gets asked by his or her boss at one time or another: “Who elected you?” ■

**David Goldston** was chief of staff of the House Committee on Science from 2001 to 2006. He can still be reached at [partyofonecolumn@gmail.com](mailto:partyofonecolumn@gmail.com).



# HOW TO SAVE A FOREST

Projects in Madagascar could provide a model for stemming deforestation. But first these efforts must deal with the poverty and political upheaval that threaten forests, reports **Anjali Nayar**.

**F**élix Ratelolahy and his team of field researchers are hiking through the dense growth of the Makira forest in northeastern Madagascar. Uapaca trees tower over them, with their spider-leg roots tall enough to walk under. Brilliant white orchids pour out of their perches in the trees. And every so often the leaves above rustle, as googly-eyed lemurs dance among the branches.

Ratelolahy, an ecologist with the Wildlife Conservation Society (WCS), and his team have spent most of the past year in this 5,200-square-kilometre forest to determine how much carbon is stored there. And today, after a three-hour hike up a precipitous slope to their first survey point, the team methodically gets to work setting up circular plots and measuring the diameter of trees (pictured above).

"We record all the numbers — the trees, the dead wood and the leaf litter," says Ratelolahy, looking up from his clipboard. "And then back in the capital, poof, the computer calculates the amount of carbon in the forest."

On these multiple-week traverses through the forest, Ratelolahy has glimpsed much of the region's endemic beauty, such as the leaf-tailed gecko and the all-white silky sifaka — a type of lemur that is one of the rarest animals in the world. But his missions have also been disturbing, he says. Over the years, Ratelolahy has watched subsistence farmers slash and burn away the margins of the forest to grow rice. And he has come across gangs pillaging the forest for rosewood, ebony and quartz. "It looks as though bombs have fallen on the place," says Ratelolahy about the ransacked areas.

**"It looks as though bombs have fallen on the place."**

— Félix Ratelolahy

Makira is on the front line of the war being waged to slow global warming. As one of Madagascar's largest forests, it stores millions of tonnes of carbon. But as in most forests in the country, that carbon is being rapidly released to the atmosphere as trees are cut down for agriculture, timber, mining and firewood. The WCS, in collaboration with the government and other organizations, is hoping to protect Makira and, at the same time, generate money

to support local communities by 'renting' the forest to rich countries.

The idea is that wealthy nations could meet their greenhouse-gas emissions targets in part by buying carbon credits from developing countries such as Madagascar. The poorer nations could earn money by keeping their forests standing, rather than cutting them down (see 'A growing market', page 28).

This strategy, known as reducing emissions from deforestation and forest degradation (REDD), is one of the topics up for discussion at the UN climate-change summit in Copenhagen this December. Countries will negotiate whether REDD should be included in the climate deal that takes over from the Kyoto Protocol when it expires in 2012.

Proponents for REDD say that this mechanism is key to cutting deforestation, which accounts for around 20% of greenhouse-gas emissions. It is also estimated that REDD could generate billions of dollars each year for forest conservation, far more than is currently spent. Hoping to cash in on the future market, projects have burgeoned around the developing

A. NAYAR



world, with those in Madagascar being some of the earliest to take shape. The projects are also helping to establish technical standards and methodologies for carbon accounting.

REDD projects must keep the promised forests standing. To succeed, this means addressing the poverty and political instability in developing countries that often lead to deforestation. These problems are particularly acute in Madagascar, where a coup earlier this year disrupted conservation efforts and raised questions about the future of REDD there.

Non-governmental organizations such as the WCS and Conservation International are working through the turmoil. But even they are worried. "We could have a very difficult time selling carbon if this political situation becomes the norm," says Lisa Gaylord, head of the WCS in Antananarivo, the country's capital. "Why would an investor want to come here?"

### Hungry for land

Madagascar is one of the wealthiest countries in terms of biodiversity, but its people are among the world's poorest. Around 85% of the population live below the World Bank's \$2-a-day poverty line and most rely heavily on the country's natural resources.

The hilly countryside is scarred by slash-and-burn agriculture, locally known as *tavy*. Once people fully exploit the fertile river valleys, they head uphill, clearing the forests to cultivate rice, the country's staple food. These rain-fed fields are harvestable for only a few seasons before productivity drops and villagers clear new land for their crops.

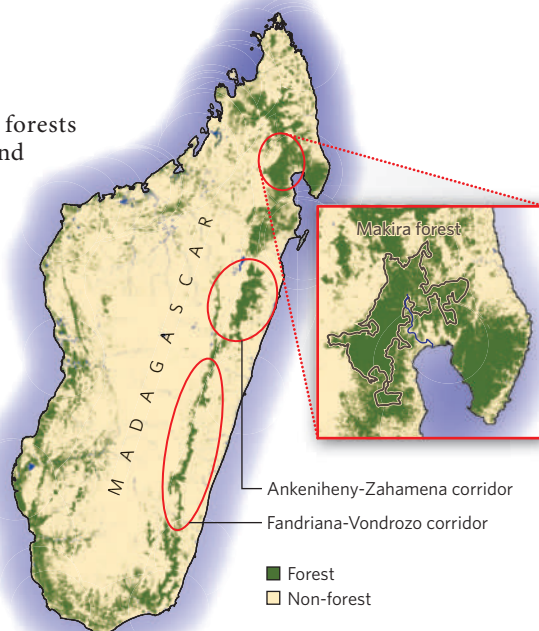
Estimates of Madagascar's original forested areas vary widely, but some studies suggest that trees once blanketed 90% or more of the island. Since aerial photographs of the

country were taken in the 1950s, forests have decreased by more than 40% and by about 2005, they covered only around 15% of the country.

Over the past two decades, deforestation has been decreasing slowly with the creation of protected areas. Grants from the World Bank and USAID helped Madagascar to become one of the first countries in Africa to develop and implement a national environmental action plan.

By 2000, the nation had protected 17,000 square kilometres of forest, mainly as national parks. But further expansion of the park system stalled because the funding stream from donors dried up. So, in 2001, the government teamed up with non-governmental organizations and started exploring the idea of selling carbon credits from their forests. "It was clear that there was a carbon market emerging and that avoided deforestation could be a very powerful way of protecting forests in Madagascar," says Frank Hawkins, vice-president for Africa and Madagascar of Conservation International in Washington DC. Hawkins was part of the team that popularized REDD on the island.

The efficacy of a REDD project depends on how much carbon the project will prevent from being released in the absence of protective measures. Calculating that number requires first measuring the current carbon content of a forest — as Ratelolahy and his crew are currently doing — and then projecting future deforestation rates with and without the project in place. The WCS is doing that by using past satellite imagery and making forecasts that account for factors such as the proximity of the



forest to roads and villages.

A study in 2004, conducted in collaboration with the non-profit organization Winrock International in Little Rock, Arkansas, estimated that the annual rate of deforestation was 0.15% in Makira. The analysis projected that the rate would rise to 0.2% a year by 2034 without any intervention. But with the REDD project in place, the deforestation rate would slow to about 0.07%. These preliminary estimates indicated that the 30-year project would avert the release of more than 9 million tonnes of CO<sub>2</sub> equivalent, similar to taking 2 million cars off the road in the United States for a year.

Madagascar's other two active REDD projects, in the Ankeniheny-Zahamena and the Fandriana-Vondrozo forest corridors, run by Conservation International, are each projected to prevent emissions of 9 million to 10 million tonnes of CO<sub>2</sub> during the same time period.

Together, at a conservative price of US\$5 per tonne, the REDD schemes in Madagascar could yield up to \$5 million a year for conservation and community development in the country, about the same as the budget for the national park system. Christopher Holmes, the technical director of the WCS's Madagascar programme, says 50% of the money from future carbon sales will go to the communities. The rest will be used to cover the costs of running, monitoring and marketing the project.

In the future, the WCS intends to pay the affected communities directly using the carbon money from Makira, but currently there is no distribution mechanism in place. In the meantime, the money will go to health and development projects aimed at reducing poverty. "People are not walking five kilometres to find forest to cut down to plant rice on a 30° slope because that's the best thing to do," says Holmes. "They are doing it because it's the only thing to do."



Félix Ratelolahy surveys rice paddies in land once covered by forest.



But setting up a REDD project in this remote part of the country is challenging, he says. Without much existing infrastructure there, the WCS has had to establish community organizations and legislative bodies that allow each community to manage their forests.

In Madagascar, the government owns the forest, so the conservation organizations are helping local communities to gain rights over the natural resources through management contracts. But progress is slow. "Some of the communities require two days of driving and three days of hiking on foot to reach," says Holmes. Since the project started in 2003, only one-quarter of Makira's 83 community organizations have signed the contracts with the government.

The contracts give communities legal access to a buffer zone surrounding a core protected area of the forest. When someone in the community wants to build a new house or dugout canoe, that person applies to the community's

organization for a permit to use the buffer zone.

Because such agreements effectively limit how much local residents can take from the forest, the WCS is trying to help communities in other ways, through projects to increase rice production and by expanding the country's eco-tourism industry into the Makira region. But for now, communities such as the remote village of Andaparaty on the eastern cusp of Makira do not see the potential benefits of carbon sales.

### Pressures on the forest

At sunrise, the riverside village is already bustling. Women are busy laying out a patchwork of woven mats topped with drying rice, beans and vanilla. Later, their meagre produce will be taken downstream to be sold.

Barnety, a man nearing 80, watches the ebb and flow of his village. He is a *tangalamena* or traditional leader in the community. When he was a boy, he says, there were only a few families in Andaparaty. Now there are hundreds of

people competing for the same land. Bit by bit, the forested slopes around the village have been stripped away to plant rice and cassava. When the Makira project reached Andaparaty in 2004, Barnety supported it. "Our lives depend on the forests," he says. "If there aren't any forests, if there isn't any land, we can't live."

But without a sustainable source of income and food, some villagers are finding it difficult to accept restrictions on their access to the forest. "People are frustrated because before the project, they were completely free to hunt, fish and cut down the forests," says Cressant Rakotomanga, president of the local community organization.

It could be years before carbon payouts come through a UN-regulated REDD system. The WCS and its donors have already spent \$1.9 million to establish the Makira REDD project, and the support going to local communities will increase when the carbon funds arrive. But for the people of Andaparaty, the support can't come soon enough. "People are wondering where the money from REDD is, or if it will ever come," says Ratelolahy.

Political unrest in Madagascar makes the future of REDD projects there even less certain. On 17 March this year, following two months of protest, the 35-year-old mayor of Antananarivo, Andry Rajoelina, took the presidency in a military-backed coup.

The ensuing political instability caused a surge in illegal wildlife trafficking, mining and logging activities, threatening the country's forests. "Because of the political instability, people feel liberated, which translates to more exploiting," says Haja Salava, the director of Masoala National Park, adjacent to Makira.

Armed gangs are ravaging the northeastern region for its valuable rosewood and ebony. The park rangers and community members who have tried to impede the illegal trade have been threatened with death, says Salava.

In the past few months, thousands of illegal loggers have been raiding his national park, Salava estimates. In his dilapidated office in the regional capital Maroantsetra, he scrolls through his monthly reports to the government. The pages are a collage of photos of men posing with freshly cut wood. The loggers aren't afraid of being caught because, despite Salava's repeated calls for assistance, the government police do not stop the trade, he says. "It's a free-for-all."

Mariot Rakotovo, who led the country's Ministry of the Environment and Forests at the time, said his department had ramped up police patrols in the region and had fined illegal loggers.

But the Missouri Botanical Garden, which has been following the illegal timber trade, estimates that more than 850 shipping

## A growing market

Many companies and individuals are going carbon neutral by purchasing credits to offset their emissions. Some of these credits pay for forest protection in tropical countries — essentially keeping the carbon locked in the living trees. Until recently there was a lack of accepted standards to guarantee that such projects result in real emissions reductions, says Toby Janson-Smith, senior director for forest carbon markets at Conservation International, which runs forest programmes around the world. "Each project did it in its own way and you had to take the word of the project that they had done a good job."

These credits are traded on

a voluntary carbon market, which has grown rapidly in the past year to more than \$700 million.

At present, forest-protection projects are not included in the carbon market used by nations to satisfy their emissions pledges under the Kyoto Protocol. This 'regulatory market' has stricter standards, and carbon on it trades for about double the price of carbon on the voluntary market.

Some nations with rainforests are pushing to include forest-protection projects in a new climate treaty, which is expected to be signed in Copenhagen in December. That would allow developing nations to

access the more lucrative carbon market for their forest programmes.

For now, the voluntary market, and the forest projects it includes, serve as a testing ground for methodologies and practices, says Janson-Smith.

For example, within the past year, organizations have developed Voluntary Carbon Standards (VCS) certification for forest-protection projects. These standards contain provisions to guard against the possibility that fires, disease or logging might destroy the protected trees.

This works by requiring that forest projects keep a buffer; they can sell only a fraction of the carbon they actually store. The buffer depends on an assessment of the project's risk. "In some cases the buffer may be as low as 20% of the carbon stocks, in other cases it may be 40–50%," says Ray Victorine, the director of the Wildlife Conservation Society's conservation finance programme. The buffer credits from the VCS-certified projects will be pooled together as an insurance policy if one or more of the projects fail.

A.N.



A. NAYAR



**Proceeds from carbon sales could help villagers from Andaparaty, on the edge of Makira forest.**

containers of rosewood have left the country since January, when the political problems began. Almost all of the exported wood was from illegal logging, says Porter Lowry of the Missouri Botanical Garden, which has conservation programmes in Madagascar.

The environment ministry did not respond to requests from *Nature* for comment. A government official, who asked not to be named, challenged the claim about such shipments. "That's a lot of wood. I don't believe it," he says.

**Wood for sale**

Other details of the timber trade are more certain. On 21 September, the government issued an inter-ministerial order allowing another wave of timber exports. The new shipments would have a combined value of tens of millions of dollars. According to the ministerial order, the shipments are designed to empty the stockpiles of timber in the country's ports and are limited in number to discourage operators from returning to the forest to cut more precious wood.

The order also stipulates that taxes on the exports will raise money for the central treasury and for forest conservation through a new fund called Action against the Degradation of the Environment and Forests (ADEF).

Jean Roger Rakotoarijaona, chair of the REDD technical committee that is putting together the country's national strategy, says legitimizing the exports of illegally felled wood

will only propagate more cutting in Madagascar. He also questions the purpose of the new forest fund. "We already have a national forest fund that needs replenishment," he says. Regarding the new fund, he says, "I'm a little worried about what this is going to be used for."

The ongoing illegal felling of trees could lead to a lot of extra carbon emissions. But even worse for Madagascar's REDD projects, the country's political instability has caused international donors to cut funding for forest and development programmes.

In March, the World Bank's Forest Carbon Partnership Facility froze a \$200,000 grant supporting Madagascar's work preparing a national REDD strategy. That delay imperils Madagascar's chances of winning a follow-up \$3.6-million grant to implement the plan, says Rakotoarijaona.

The environment and forests ministry also lost about 95% of its funding, which had come from international donors, says Rakotovo.

The cuts have severely impeded the ministry's ability to manage and patrol the country's forests, he says.

Investments in the country's carbon-credit projects have also stalled. The World Bank's BioCarbon Fund put

a hold on its initial payment to buy offsets for 1.5 million tonnes of CO<sub>2</sub> emissions from one of Conservation International's REDD and reforestation projects in eastern Madagascar.

The loss of funding for development projects that provide rural communities with alternative livelihoods also means more people are returning to slash-and-burn agriculture. "The short-term pillaging of forests is a problem, but

slash-and-burn agriculture is the worst poverty trap," says Hawkins. "In Madagascar, it leads to the permanent loss of forests — you can't dig yourself out of that."

In the past few months, pressure from the international community forced Rajoelina into power-sharing talks with the country's three former governments. The aim is to bring about elections before the end of next year. But even if the political situation stabilizes and investors come back, there is a long road ahead to get REDD working in the towns bordering Makira.

The WCS is running workshops to teach villagers how to increase their crop production. But few people have adopted the improvements, says Jean Jaonary, the local community organization president in the village of Ambodivoahangy, on the northeast side of Makira. "Using the new method, my rice production has doubled," he says, while wading through the neat rows of his lime-green fields. "I don't know why other people haven't caught on."

As Jaonary walks through Ambodivoahangy's rice fields with Ratelolahy, women and men are busy with work, their straw hats popping in and out of the greenery. Ratelolahy says the land is all the rural populations have, and they are weary of new ways of farming and efforts to keep them out of the forest. He has faith that REDD can help these communities, but it will take time to convince each of the villagers to change. Looking towards the future, Ratelolahy summons up an old Malagasy saying: "Cows don't all wake up at the same time."

**Anjali Nayar is an International Development Research Centre Fellow at *Nature*.**

**See Editorial, page 11. Watch an audio slideshow at [go.nature.com/VGGayN](http://go.nature.com/VGGayN)**

**"Because of the instability people feel more liberated, which translates into more exploiting."**

— Haja Salava

A. NAYAR



# Reflecting the past

Unsatisfied with merely halting environmental destruction, some conservationists are trying to reconstruct ecosystems of the past. **Emma Marris** travels back in time with the rewilders.



**F**enced off from the modern Dutch countryside is a scene that looks more like a diorama in a natural-history museum than a typical central-European park: a plain is dotted with wild herds of large mammals. Empty of humans and haunted by eagles, it is a vision of a distant past on some of the newest land on Earth — the reserve was only reclaimed from the sea in 1968. It's called the Oostvaardersplassen, and the visionary behind it is Frans Vera, a tall, greying government scientist from Staatsbosbeheer, the organization responsible for overseeing Dutch nature reserves.

Vera designed the 6,000-hectare reserve to replicate Europe's prehistoric past. That has meant 'rewilding' the area, populating it with the kinds of creature that lived there many thousands of years ago. That some of these creatures — including wild horses called tarpans, and aurochs, wild ancestors of modern cattle — are extinct, was not enough to stop Vera. In the 1980s, he went shopping for substitutes. He acquired Konik horses, believed to be descendants of wild Tarpans, from Poland. For the aurochs, he substituted Heck cattle, a line developed by two German brothers in the early twentieth century from a number of cattle breeds with the intention of mimicking auroch features. The herds number in the hundreds and graze alongside a red-deer population of about 2,000. Next, Vera would love to get his hands on some European bison and boar.

Vera's reasons for developing this eclectic menagerie might not be immediately clear. It isn't a tourist attraction; few people are granted entry, although viewing points for bird watchers

dot the periphery. It has, however, succeeded as a conservation area; several bird species rare to Western Europe, such as the white-tailed eagle, have moved in. Vera says that it is also a large science experiment, designed to test his theories about how European landscapes used to look. But he has struggled to keep the reserve open and lacks the funding for graduate students. The project generates very little systematic data or scientific papers, adding to its mystique.

## Pleistocene parks

Vera's isn't the only rewilding project. Schemes in locales as diverse as New Zealand, Saudi Arabia and the Russian Far East aim to do more than hold the line against further environmental destruction (see 'Lost landscapes'). They are attempting to recreate the ecological workings of previous eras, often those of the Pleistocene from 2 million to 10,000 years ago. Advocates for Pleistocene rewilding would restore, if they could, the age when nature lived wild and large, when mastodons, sloths and three-tonne wombats heaved their bulk around Earth, and predators were big, fast and ubiquitous. Because many of the fauna in these systems are now extinct, the schemes often turn to proxies to fill abandoned roles: grazing, browsing, defecating and culling the herd.

Critics, however, say that the projects are more sentiment than science. Dustin Rubenstein, an ecologist at Columbia University in New York, argues that placing proxy animals in a modern landscape is not the same as turning back the clock, and could spell trouble<sup>1</sup>. These ecosystems have changed and existing species

have evolved in the thousands of years since megafauna extinctions. Attempting to fill gaps that closed long ago with proxy animals can generate unpredictable results. "It is the difference between the known and the unknown," he says. Proxy animals, he suggests, could become invasive pests, or escape their parks and cause trouble with local landowners, who would then turn against the conservationists. Moreover, Rubenstein worries about how the projects in Europe and elsewhere are being run and says that information about them is not being disseminated. "We're not seeing the results in the peer-reviewed literature." Without anything being published, he says, rewilders can't dispel worries about the consequences of such proxy introductions.

Vera says that the complaint has merit, but that the project doesn't have enough money or personnel to monitor and publish on the wildlife in the reserve. William Sutherland, an ecologist at the University of Cambridge, UK, with a special interest in documenting conservation experiments, says that it's a common, frustrating occurrence. "It is as if you can only find out about details of a cancer treatment by going round the wards with the doctor," he says. Sutherland says that Vera and other rewilders — along with many other conservation biologists — are "too busy conserving to have the time to monitor and publish".

The only way to learn about the Oostvaardersplassen is to go there. Nestled between windmills, small dairy farms, a dyke and the city of Almere, it is, above all, totally unexpected: a Serengeti of sorts with vast, grassy

R. SMIT



plains where wild horses, cattle and red deer move in massive herds. Two white-tailed eagles perch near their nest, an enormous tree house of bone-white branches. Spoonbills, greylag geese, ibises and many other European bird species converge on the marshes. From the taller grass, a white and orange mist of butterflies explodes in front of Vera's van. And on one edge of the park, an abandoned calf — its eyes glazed, its knees wobbly — faces death.

Vera watches the calf for a few moments, then reaches for his radio and calls one of the reserve's staff members, who will come and kill it. Although the park is supposed to support nature in its raw form, the Dutch public considers letting animals starve to death too cruel to stomach, so Vera has worked out an agreement with the government that provides swift dispatch to sick and starving animals. Every year, 10–20% of the large herbivores in the park die from natural causes or are killed by humans. But their corpses in turn support other animals. Vera was perhaps the only person not surprised to see a Eurasian black vulture, long absent from the area, appear in the park — although its tenure was short. It strayed outside the fence and was run over by a train.

The Heck cattle at the Oostvaardersplassen could, at a pinch, be mistaken for the aurochs painted in famous Palaeolithic caves at Lascaux in southwestern France. Large-bodied and wide-horned, they come in several colours, from black to brown to beige. "For me, the point was that they not look like dairy cattle," says Vera. He wants people to see them as wild animals, and — apart from the compromise reached on mercy killings — they live a wild life, without veterinary care, supplemental feeding or even ear tags. They do what comes naturally, and in so doing they are keeping the reserve grazed down and open, turning the earth with their massive hooves and, in their wallows, creating depressions that fill with water and become mini-ecosystems.

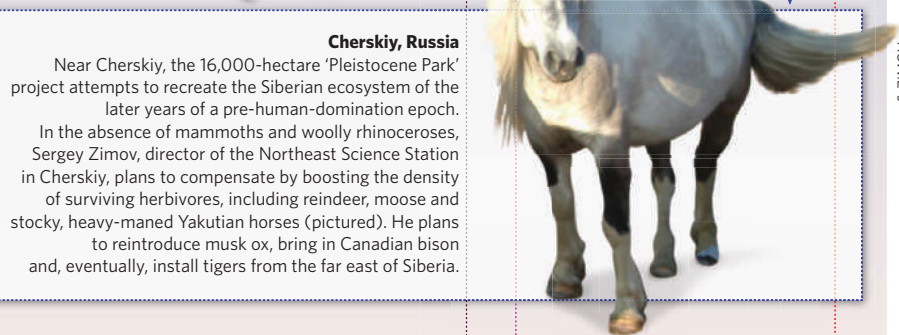
### Reshaping landscapes

Vera hopes that the Oostvaardersplassen will help to answer a couple of scientific questions: can grazers drive ecosystem processes, and will the resultant landscapes be patchy or densely forested? Vera's vision of Europe's past is of a patchy landscape, what he calls a "park-like landscape" in which any given area cycles between open grassland maintained by herbivores, thorny shrubs that sneak in where herbivores are caught napping and tree-dominated groves that grow in the shelter of the herbivore-proof shrubs.

The sequence of events could explain why oak seedlings don't grow up to be adult oaks in the shade of other trees, even though giant oaks

## LOST LANDSCAPES

Efforts to rewild from around the world.



**Cherskiy, Russia**

Near Cherskiy, the 16,000-hectare 'Pleistocene Park' project attempts to recreate the Siberian ecosystem of the later years of a pre-human-domination epoch. In the absence of mammoths and woolly rhinoceroses, Sergey Zimov, director of the Northeast Science Station in Cherskiy, plans to compensate by boosting the density of surviving herbivores, including reindeer, moose and stocky, heavy-maned Yakutian horses (pictured). He plans to reintroduce musk ox, bring in Canadian bison and, eventually, install tigers from the far east of Siberia.

S. ZIMOV



**Mahazat as-Sayd, Saudi Arabia**

In central Saudi Arabia, the Mahazat as-Sayd is fenced off from grazing livestock. Inside its more than 200,000 hectares, reintroduced houbara bustards, reem gazelle and Arabian oryx (pictured) — a spiral-horned ungulate that went extinct in the wild in the 1970s — mingle with the African red-necked ostrich, a proxy for the extinct Arabian subspecies that once strutted here<sup>3</sup>. Native vegetation and birds largely complete the picture of the Arabian Peninsula as it might have looked 2,000 years ago. But wolf and cheetah predators are absent.


O. COUPPEY/NWRC PHOTO LIBRARY



**Wellington, New Zealand**

Zealandia (formerly the Karori Wildlife Sanctuary), 225 hectares in the heart of Wellington, allows urbanites to experience the country more or less as it was before people stepped ashore. A mammal-proof fence protects many rare bird species, including some, such as the little-spotted kiwi (pictured), that had not been seen on the mainland for many years. The scientists behind the reserve are toying with adding a few proxies as well — having the Australian owl-nightjar play the extinct New Zealand owl-nightjar, for example. Scientific data are being collected in reams, but peer-reviewed publication lags, thanks to a dearth of time and money.

ZEALANDIA/KARORI SANCTUARY TRUST



**Lake Pape, Latvia**

On 5,700 hectares near Lake Pape, Heck cattle, Konik horses and bison (pictured) roam in a project inspired by the Oostvaardersplassen in the Netherlands. Ecology consultant Joep van de Vlasakker says that the area initially had more promise for growth. But Latvia's entry into the European Union brought subsidies for small farmers, many of whom have now set up shop around the reserve. Van de Vlasakker says the soil is poor and the farmers are unsuccessful. "They mow and plough, but they don't harvest," he says. Meanwhile tourism could benefit from the ecological oddity: "The project could be very influential for local economy."

J. VAN DE VLASAKKER/FLAXFIELD NATURE CONSULTANCY



Heck cattle can act as proxies for the extinct aurochs.

are found in all of the continent's ancient forests. Vera believes that most old oaks grew up in open areas and then saw the forest fill in around them. His is the minority view. Most scientists think a closed forest covered the continent. So far, the Oostvaardersplassen has shown that a high density of grazers can certainly affect the landscape: they have largely mowed it clean. Vera suspects that thorny shrubs will establish themselves and act as nurseries for tree seedlings, even if it doesn't happen in his lifetime.

### Rough approximations

Vera's experimental set-up does have major limitations. As he and other designers readily admit, every rewilding project will be only an approximation of a past ecosystem. With taxon substitutions and incomplete sets of plant and animal functional groups, many of these systems perhaps evoke the past more than they replicate it.

The Oostvaardersplassen, for example, contains none of its lost predators, such as bears or wolves, yet other reintroduction experiments have shown that they can alter the entire ecosystem. When wolves were reintroduced to Yellowstone National Park in Wyoming, for example, the elk soon learned to feed only in areas where they could see wolves if they approached. As a result, areas with bad sight lines were soon thick with willow and cottonwood seedlings<sup>2</sup>. Perhaps this 'ecology of fear' would promote the growth of the thorny shrubs that Vera eagerly awaits. However, weary from battles over the management of the Oostvaardersplassen, he says he is in no hurry to step into the political minefield of wolf reintroduction in a country where livestock owners are a powerful political force. Instead, he says, he will wait for wolves from increasingly popular reintroduction programmes elsewhere in Europe to converge on the Netherlands. Wolves are notoriously good at covering ground, even in population-dense

areas such as northern Europe. "They will come, whether we like it or not," he says.

Josh Donlan, director of the conservation consultancy Advanced Conservation Strategies in Midway, Utah, and a champion of Pleistocene rewilding projects, says that for some projects the most important motivations are not science- but conservation-oriented. Others, such as a tortoise proxy programme in the Mascarene Islands near Madagascar are "completely hypothesis driven", says Donlan, and being carefully documented. These reintroductions offer a huge opportunity to do science, he says, not only to learn about the ecology of the past, but also to feed back into the projects, "which are going to need all the information they can to be successful".

Donlan and others have proposed that Pleistocene reserves be created in North America, where African and Asian animals could fulfil the roles of the large mammals that went extinct on the continent about 13,000 years ago. The proposal would have two conservation goals. Lineages of species that have been driven to extinction would have a chance to return to North American ecosystems and start to evolve along their own unique path, and species that would themselves be saved from extinction in their home countries would resuscitate extinct processes in North America. Elephants, for example might eat large fruit such as the osage orange, dispersing their seeds naturally as mastodons and other elephant relatives were thought to do thousands of years ago. This vision of vast areas populated with camels, wild asses and lions has caused a major stir. It has generated strong responses from the public, both for and against, with many dismissing the whole idea as insane.

But Donlan says that many people are coming around to the concept. "I think that it is certainly gaining traction in the public view,"

he says. "Conservationists are starting to rally around this idea, almost certainly partly because it is proactive." That is, it is something that conservationists can do beyond trying to prevent development and negative change — a strategy that Donlan considers pessimistic and mostly ineffective. Rubenstein has problems with Pleistocene rewilding as it stands, but gives credit to Donlan's desire for change.

For Stuart Pimm, a conservation biologist at Duke University in Durham, North Carolina, and one of the few scientists to have set foot in the Oostvaardersplassen, the re-appearance of white-tailed eagles, breeding barnacle geese and at least one black vulture is proof that the reserve is achieving its conservation goals. "They've got it right and a lot of things have come back as a

consequence," he says. As for the scientific goals, he urges a broader view. Sure, there are no control areas or replications. But at this scale, how could there be? "The idea that we have to do neatly replicated experimental design with little squares doesn't capture the

whole range of what science does," he says.

Vera is approaching retirement age now, and looking for a successor to take over the job he's been doing, under various titles, for 30 years. For his retirement, who knows. Perhaps he'll consult on rewilding projects elsewhere. "I would love to find out what an elephant eats in the temperate regions," he says. ■

**Emma Marris writes for *Nature* from Columbia, Missouri.**

1. Rubenstein, D. R., Rubenstein, D. I., Sherman, P. W. & Gavin, T. A. *Biol. Conserv.* **132**, 232–238 (2006).
2. Ripple, W. J. & Beschta, R. L. *BioScience* **54**, 755–766 (2004).
3. Seddon, P. J. & Soorae, P. S. *Conserv. Biol.* **13**, 177–184 (1999).

**See [go.nature.com/kT4vnz](http://go.nature.com/kT4vnz) for a slideshow of images from the Oostvaardersplassen and an interview with the author.**

**"The wolves will come, whether we like it or not."**  
— Frans Vera



# CORRESPONDENCE

## Sharing: lessons from natural history's success story

**SIR** — The success of data sharing among natural-history collections may alleviate the fears about capacity and cooperation expressed in your special issue (*Nature* **461**, 145, 160–163 and 168–173; 2009). Our social and information-technology (IT) infrastructure provides open access to millions of records from hundreds of repositories, thanks to broad participation and funding from the US National Science Foundation. However, this success story also exposes some new challenges.

Key to this success has been the development of a distributed publishing system that conserves full rights of contributors to data and access. Data are curated at source and then made available to the community for use and improvement. The Mammal Networked Information System (MaNIS) was established in 2001 to provide hardware, data standards, transmission mechanisms and stakeholder support for open access to specimen data.

Since then, more thematic, taxonomic and regional networks have been created. Each of these feeds into a growing biodiversity-informatics community, including the Global Biodiversity Information Facility, which serves users with up to 189 million plant and animal records. The vertebrate-based biodiversity networks — including MaNIS, HerpNet, ORNIS and FishNet 2 — serve some 4% of their combined holdings each day to users hungry for these data.

Sustaining these resources is difficult. Growth has led to problems with scalability and sustainability, including difficulties in keeping resources running, slow provider response times and complicated installations and maintenance. The National Biological Information Infrastructure has provided

support, and the vertebrate networks are consolidating into a platform called VertNet. In order to reduce IT costs, VertNet will move from institutional servers to a cloud computing platform, providing nearly unlimited room for growth.

But solving technological challenges is not enough. Our success has depended on strong engagement with our contributor and user community. Capacity building in biodiversity informatics is especially important. Success requires willing participation, robust technology choices and a commitment to engage fully with the communities these repositories will serve.

**Robert Guralnick** University of Colorado Museum of Natural History, University of Colorado–Boulder, Boulder, Colorado 80309, USA  
e-mail: robert.guralnick@colorado.edu  
**Heather Constable, John Wiczorek, Craig Moritz** Museum of Vertebrate Zoology, University of California–Berkeley, Berkeley, California 94720, USA  
**A. Townsend Peterson** Natural History Museum and Biodiversity Research Center, University of Kansas, Lawrence, Kansas 66045, USA

## No final answers yet on sex determination in birds

**SIR** — Readers of J. A. M. Graves's News & Views article 'Sex determination: Birds do it with a Z gene' (*Nature* **461**, 177–178; 2009) may be left with the impression that a particular gene, *DMRT1* on the Z chromosome, is the master determinant of sex in birds. However, the case is far from closed.

Graves discusses the findings of Craig A. Smith and colleagues (*Nature* **461**, 267–271; 2009). These data support the view that *DMRT1* is required for testis development in chickens. But as the authors state, "Although Z-linked *DMRT1* is required for testis development, it is possible that another Z-linked gene lies

upstream of this gene in the avian male-determining pathway ... It is also possible that a female determinant lies on the avian W sex chromosome".

It may turn out that *DMRT1* does determine sex in birds. But as yet other possibilities remain open.

**Asato Kuroiwa** Laboratory of Animal Cytogenetics, Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan  
e-mail: asatok@ees.hokudai.ac.jp

## Toxicity testing by category for 30,000 chemicals?

**SIR** — Thomas Hartung and Costanza Rovinda in their Opinion article (*Nature* **460**, 1080–1081; 2009) argue that the European Chemicals Agency has underestimated the impact of European Union (EU) legislation on animal testing by six times, and that some toxicity tests should be suspended. Unfortunately, pragmatic solutions for the implementation of reduced animal testing under the EU's REACH legislation — for registration, evaluation, authorization and restriction of chemicals — have not yet been formulated in a clear operational manner.

On the basis of the experience in the US High Production Volume Challenge Program, where non-testing approaches have been applied successfully, the goals for REACH can be formulated in a 'smart' manner. In this US programme, around 81% of the chemicals were included in a chemical category and new testing was proposed for fewer than 10% of the human-health and ecotoxicity end-point data needed (K. Van Leeuwen *et al.* SAR QSAR Environ. Res. **20**, 207–220; 2009).

If the United States can waive testing for most of its industrial chemicals, Europe can strive for at least 50% reduction in animal testing over the next five years by applying a similar approach. This

will allow us to move away from the slow, traditional, chemical-by-chemical procedures.

Achieving regulatory acceptance for such approaches in Europe is the main challenge. REACH has to assess at least 30,000 chemicals in the next 11 years, an increase of a factor of 300 compared with the past 14 years (G. Schaafsma *et al.* Reg. Toxicol. Pharmacol. **53**, 70–80; 2009). Hence, category approaches are the way forward.

**Kees van Leeuwen, Gerwin Schaafsma** TNO Quality for Life, PO Box 360, 3700 AJ, Zeist, The Netherlands  
e-mail: k.vanleeuwen@tno.nl

## Authors beware, and protect your online identity

**SIR** — Goudarz Molaei is right to express concern in his Correspondence about simultaneous submission of manuscripts to different journals (*Nature* **461**, 723; 2009). As a professional journal editor with more than 20 years' experience, I would like to highlight here a worrying new problem I recently encountered: duplicate submission arising from author impersonation.

Unfortunately, online submission and review systems inadvertently encourage this unwelcome activity. For example, a co-author or colleague may be given the corresponding author's account password in order to submit his or her manuscripts — perhaps because of the corresponding author's lack of time or unfamiliarity with file creation and uploading. These people are then able to change the author's accounts, including the passwords, and submit manuscripts in that person's name without their knowledge.

So, authors, be wary of who has access to your account. Keep a check on what's happening and change your password after files have been submitted.



**"Newton prosecuted with relish the counterfeiters partly responsible for the disarray of the country's currency." Robert Iliffe, page 39**

ILLUSTRATION BY DAVID PARKINS

Better still, contact the editorial office if you need any help with file uploading.

**Irene Hames** *The Plant Journal*,  
Department of Biology (Area 2),  
University of York, York YO10 5DD, UK  
e-mail: imh5@york.ac.uk

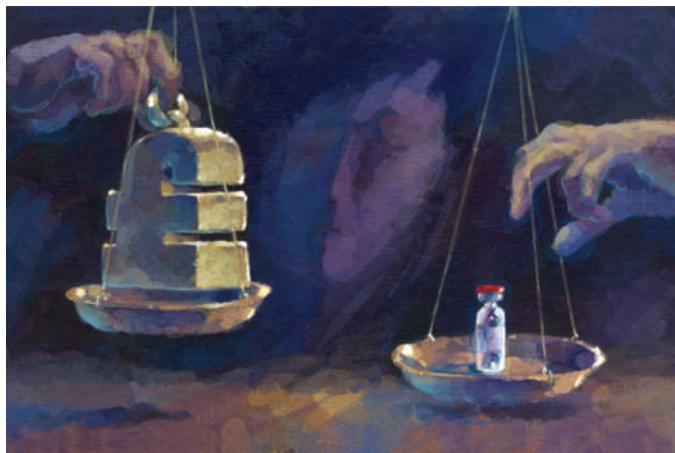
## NICE should value real experiences over hypothetical opinions

**SIR** — The UK National Institute for Health and Clinical Excellence (NICE) and other agencies for assessing health technology around the world are facing up to the challenges of rationing in a systematic and transparent way. But consideration of two factors could improve their decision-making.

First, the quality-of-life evaluation mentioned in your News Feature (*Nature* **461**, 336–339; 2009) needs more thought. NICE would achieve more if it valued health interventions according to the real suffering of patients, rather than on the basis of the hypothetical preferences of the public. There is evidence showing that the public are often prepared to sacrifice more life years than patients might be.

Also, public and patient preferences can misrepresent the impact of a particular state of health on our experiences (P. Dolan and D. Kahneman *Econ. J.* **118**, 215–234; 2008). For example, we may imagine physical pain to be more severe than depression, but depression can make us feel worse and so we evaluate our lives less favourably.

Second, NICE should not raise the cost per quality-adjusted life year (QALY) threshold for some conditions, such as the end of life, until there is good evidence for doing so. The threshold varies across different conditions. From an implicit default position where all QALYs are treated equally, NICE can now give greater weight to QALYs at the later stages of a terminal disease. NICE justifies this position as being in



accordance with the views of the general public — yet the evidence in this regard is actually quite weak.

There is some support from NICE's Citizens' Council for spending more on end-of-life care, but this preference has not been elicited in the context of what people would give up for it. In a choice between prioritizing end of life and reducing inequalities in lifetime health, it is likely that the general public would choose the latter (see [go.nature.com/QgnrFX](http://go.nature.com/QgnrFX)).

**Paul Dolan** *Imperial College Business School, Tanaka Building, Exhibition Road, London SW7 2AZ, UK*  
e-mail: paul.dolan@imperial.ac.uk

## Eastern European science stuck in an outmoded system

**SIR** — The problem that Jan Konvalinka and colleagues identify in their Correspondence — that the Czech bibliometric system for assessing grant proposals encourages mediocre research (*Nature* **460**, 1079; 2009) — is widespread in eastern Europe. Some policy-makers are upholding the old systems for allocating funds in public universities and research institutes, rather than using them to promote the best science.

In Slovenia, a bibliometric system introduced a few years

ago provides the main criteria for awarding research grants by national government, often overriding peer-reviewed evaluation. It divides scientific journals into categories that disregard impact factors, on the grounds that journals vary in quality for different fields. Although there may be some justification for the thinking behind this, the system in practice favours researchers who publish in low-impact journals over those who are struggling to do internationally competitive science and to publish in the best journals.

For example, a Slovenian endocrinologist might choose to publish a basic animal study in a prestigious endocrinology journal or in a much lower-impact agricultural journal. If the latter journal is rated at a higher position among agricultural journals than the former is in the field of endocrinology, the researcher will receive more points for publishing in the low-impact agricultural journal.

Likewise, a review article in a journal with no impact factor (but included in SCI, the science citation index) could be worth as many as 20 Slovenian bibliometric points, whereas one paper that was recently published in *Nature* — the result of years of work — gained its authors just 18.75 points each.

With the exception of the former East Germany, many universities and science policies

in central and eastern European countries did not go through a transition period when the economy changed. Respected professors who grew up and were educated under a very different system are still the principal policy-makers and still represent the majority of grant holders, even though they have not necessarily proved themselves in internationally competitive science.

**Gregor Majdič** *Center for Animal Genomics, Veterinary Faculty, University of Ljubljana, Gerbiceva 60, SI-1000 Ljubljana, Slovenia*  
e-mail: gregor.majdic@vf.uni-lj.si

## Research into group differences isn't wrong, just pointless

**SIR** — Bruce Lahn and Lanny Ebenstein (*Nature* **461**, 726–728; 2009) state "Some scholars have even called for a halt to research into [genetically based group differences] because of potential misuse of information", referring to a Commentary of mine on race, gender and intelligence (*Nature* **457**, 786–788; 2009).

This is a misinterpretation of my argument. My view is as I wrote in that Commentary: "The problem is not that knowledge of such group intelligence differences [between black and white, men and women] is too dangerous, but rather that there is no valid knowledge to be found in this area at all."

Lahn and Ebenstein disagree with my contention, which is of course their privilege. But they are not correct to conclude that I believe research in this or any area should be censored.

**Steven Rose** *Department of Life Sciences, The Open University, Milton Keynes MK7 6AA, UK*  
e-mail: s.p.rose@open.ac.uk

Contributions may be sent to [correspondence@nature.com](mailto:correspondence@nature.com) and will be edited before publication. Please see the Guide to Authors at [go.nature.com/cMCHno](http://go.nature.com/cMCHno).

## OPINION

## Global Darwin: Contempt for competition

Darwin's idea of the 'struggle for existence' struck a chord with his fellow countrymen. But Russians rejected the alien metaphor, says **Daniel Todes**, in the second of four weekly pieces on reactions to evolutionary theory.

In *On the Origin of Species*, Charles Darwin acknowledged his intellectual debt to the Reverend Thomas Robert Malthus. That debt had radically different consequences for his British and Russian readers.



Darwin200

In *An Essay on the Principle of Population, as it Affects the Future Improvement of Society* (1798), Malthus argued against believers in social progress by citing an inexorable natural law: population tends to increase geometrically and food supply only arithmetically. These imbalanced progressions lead to a "struggle for existence" in which the winners prospered and the losers suffered privation and premature death. Nature itself decreed that human misery was inevitable.

By Darwin's day, Malthus's theory had entered the mainstream of British thought. Pondering possible mechanisms of evolution in 1838, the 29-year-old Darwin picked up Malthus's essay. Never a full-throated Malthusian in his political attitudes, he nevertheless adapted Malthus's idea to his science. "As more individuals are produced than can possibly survive," he explained in *On the Origin of Species* (1859), "there must in every case be a struggle for existence, either one individual with another of the same species, or with the individuals of distinct species, or with the physical conditions of life. It is the doctrine of Malthus applied with manifold force to the whole animal and vegetable kingdoms."

Darwin recognized that he was using the term "struggle for existence in a large and metaphorical sense" to encompass a variety of natural relations that one wouldn't necessarily conceive of as a battle: not just two dogs fighting over a scrap of food, but also a plant seeking moisture in the desert, or the dependence of one being on another.

For Darwin and other leading British evolutionists, this appealed to common sense. Living on a crowded island with a capitalist economy and highly individualist culture, struggle for existence did not seem a metaphor at all, but, rather, a simple and eloquent description of nature and society.

Russians, however, lived in a very different land. Their own cultural values and experiences

would lead them to reject Darwin's Malthusian metaphor. This in turn affected a wide range of research — from studies of the mutual aid among migrating fish to a Nobel prizewinning theory of inflammation and immunity — and echoed well into

the twentieth century, perhaps even playing a part in the enthronement of Lysenkoism. This Russian response provides a striking example of the way in which metaphors — and the experiences and cultural traditions that they capture — shape scientific thought.

The experiences of leading Russian naturalists were in many ways opposite to those of Darwin and his fellow proposer of evolution by natural selection, Alfred Russel Wallace. The two men shared seminal field experiences in densely populated tropical environments. The contest between organisms seemed obvious there. Most Russian naturalists, by contrast, investigated a vast under-populated continental plain. For them, nature was not an "entangled bank" — the image Darwin took from the Brazilian jungle. It was a largely empty Siberian expanse in which overpopulation was rare and only the struggle of organisms against a harsh environment was dramatic.

### Cultural divide

Russia's economy, political structure and culture also contrasted sharply with those in the United Kingdom. Capitalism was only weakly developed and political supporters of the two most important classes, rich landlords and peasants, spoke the language of communalism — stressing not individual initiative and struggle, but the importance of cooperation within social groups and the virtues of social harmony. Russian political commentators of the left, right and centre reviled Malthus as an apologist for predatory capitalism and soulless individualism.

The cultural gulf between the two lands was captured by demographer and biologist Nikolai Danilevskii's summary of the British character in his book *Russia and Europe* (1869). The typical Englishman, he wrote, "accepts [struggle] with all its consequences, demands it as his right, tolerates no limits upon it". In his two volumes on Darwinism (1885, 1889), he catalogued

the lengths to which the English went to indulge their passion for individualistic conflict. They boxed one-on-one (not in groups, as Russians liked to spar), founded debating societies for the "struggle of opinions", and even established mountain-climbing clubs, not for scholarly purposes, "but solely to allow oneself the satisfaction of overcoming difficulties and dangers ... in competition with others".

Small wonder, then, that few Russians shared Darwin and Wallace's respect for Malthus, and that many saw the struggle for existence as an infusion of the British enthusiasm for individualistic competition into natural science. Darwin's theory, as Danilevskii put it, was "a purely English doctrine".

Most Russian naturalists, many of whom were evolutionists before 1859, shared that view. Yet they also admired Darwin and didn't think his association with Malthus justified complete rejection of his theory. Their common response was to break down Darwin's Malthusian metaphor into its component parts, to explore their relationship and relative importance in nature and to conclude that he had greatly exaggerated the role of the two parts most closely associated with Malthus: overpopulation as the generator of conflict, and intraspecific competition as its result.

This common response defined a general direction, but individual scientists took different paths. Russia's leading botanist, Andrei Beketov, concluded that intraspecific struggle was a minor note within the general "harmony of nature". Devaluing natural selection, he reaffirmed his long-standing view that evolution resulted chiefly from the direct action of the environment on organisms. Botanical geographer Sergei Korzhinskii was led to his 'theory of heterogenesis' — the idea that mutations create large, step changes that could yield new species in a single move. This theory, he emphasized, offered the great advantage of denying any creative evolutionary role to the struggle for existence, which he thought merely pruned the rich tree of nature. Zoologist Ilya Mechnikov emphasized interspecific struggle. This proved crucial to his development of the 'phagocytic theory of inflammation and immunity', for which he received a share of a Nobel prize in 1908.

The critique of Darwin's metaphor led many Russian naturalists to the theory of mutual aid, which emphasized the importance of

**"In the Siberian expanse, only the struggle of organisms against a harsh environment was dramatic."**





ILLUSTRATION BY G. LAM

cooperation. Darwin too had called attention to such cooperation, but the theory of mutual aid went further. It held that the central aspect of the struggle for existence is an organism's struggle with abiotic conditions, that organisms join forces in this struggle, that such mutual aid is favoured by natural selection, and that cooperation so vitiated intraspecific competition as to render it unimportant in the origin of new species. Often voiced in the 1860s and 1870s by lay intellectuals and scientists of every political stripe, this view was first systematized by St Petersburg University's ichthyologist Karl Kessler, whose oral presentation *On the Law of Mutual Aid* (1879) transformed this widespread sentiment into a staple of Russian evolutionary thought.

### Anarchistic association

Westerners, however, soon came to associate this view with one of Kessler's admirers, the exiled anarchist prince Peter Kropotkin. In a mirror image of the Russian response to Darwin's invocation of Malthus, western Europeans often dismissed the theory of mutual aid as a simplistic expression of Kropotkin's anarchism.

Yet Kropotkin's critique of Darwin's Malthusianism had originated in 1862–67, long before he became a committed anarchist. He had travelled through Siberia with a series of military and commercial expeditions, traversing more than 80,000 kilometres in the same role of gentleman-observer that had taken Darwin, decades earlier, to the tropics. Already an evolutionist, Kropotkin read *Origin* in the Siberian wilderness, and found the emphasis on overpopulation and intraspecific competition unconvincing. As an exile in England years later,

an appalled Kropotkin read Huxley's "atrocious article" on *The Struggle for Existence in Human Society* (1888). His responses, brought together in *Mutual Aid: A Factor of Evolution* (1902), reflected the basic logic of the Russian national style, just as Huxley's essay reflected that of his own homeland.

The struggle for existence remained a preoccupation for Russian evolutionists well into the 1920s and 1930s. Among them was Georgii Gause, who developed the 'competitive exclusion principle' (which held that no two species could share the same ecological niche in a stable environment). His laboratory experiments and mathematical analyses confirmed the importance of intraspecific competition, contrary to the traditional Russian consensus.

In 1948, Joseph Stalin himself encouraged Trofim Lysenko to add an extensive critique of Darwin's "Malthusian error" to Lysenko's landmark speech about his own 'creative' Darwinism. As a young revolutionary at the turn of the century, Stalin had read Darwin and taken an interest in evolutionary theory. Lysenko's doctrine, which was forcibly imposed on Soviet biology from 1948 to 1964 by Stalin and his minions, endorsed the Lamarckian inheritance of acquired characteristics, rejected the gene as a material unit of heredity, and denied the evolutionary role of overpopulation and intraspecific competition. The long-standing Russian critique of Darwin's Malthusianism did not cause Lysenkoism, but it seems possible that, by influencing Stalin, it contributed to this tragedy.

A different metaphor caused Darwin problems in his native land. Wallace remarked, in his article *Mr. Darwin's Metaphors Liable to*

*Misconception* (1868), that the Malthusian progressions and struggle for existence were self-evident "facts". Yet because natural selection seemed to personify a perceptive and forward-thinking selector, or god, he urged Darwin to replace the term with "survival of the fittest".

Darwin, however, had brushed him off. "Every one knows what is meant and is implied by such metaphorical expressions," he had demurred. "And they are almost necessary for brevity."

On this point Darwin was surely mistaken. Metaphors are brief, but they are fruitful and powerful precisely because they are not clear. They propose open-ended associations that acquire specific meaning only in the mind of individuals who consider for themselves, based on their experiences, how precisely existence is a 'struggle', an animal is a 'machine' or DNA a 'code'. Those associations and meanings often have a cultural component.

Researchers bring their life experiences and culture with them into the field and laboratory, and in the course of their investigations actively originate, interpret, develop and reject metaphorical pathways. As is shown by the reception of Darwin's theory in Russia, the deployment and criticism of metaphors are part of the ineffably human process by which scientists mobilize their experiences and values to explore the infinite complexity of nature. ■

**Daniel Todes** is at the Institute of the History of Medicine at Johns Hopkins University, 1900 East Monument Street, Baltimore, Maryland 21205, USA. He is currently completing a biography of Ivan Pavlov. e-mail: dtodes@jhmi.edu

For further reading, see [go.nature.com/c8ABz5](http://go.nature.com/c8ABz5)

For more on Darwin see [www.nature.com/darwin](http://www.nature.com/darwin)



## BOOKS &amp; ARTS

## Amphibian mystery misread

A book blaming a fungus for the disappearance of amphibians from wild places wrongly downplays the role of environmental change, warn **Alan Pounds** and **Karen Masters**.

**Extinction in Our Times: Global Amphibian Decline**

by James P. Collins and Martha L. Crump  
Oxford University Press: 2009. 304 pp.  
\$29.95

Across the globe, frogs, toads and salamanders are disappearing, even in protected habitats. In *Extinction in Our Times*, James Collins and Martha Crump try to reassure us that these vanishing creatures are not warning of large-scale environmental deterioration like canaries in a coal mine, but are simply “telling us that they themselves are in trouble”.

The cause of amphibian declines in wild places, the authors assert, is often a single agent: the chytrid fungus *Batrachochytrium dendrobatidis*, which can produce a fatal skin disease in these animals. According to this view, the mass die-offs result from movement of the fungus between continents and across landscapes. However, this ‘lone killer’ hypothesis wrongly downplays the role of environmental change.

Collins and Crump see evidence for a lone-killer chytrid in reports that a wave of amphibian extinction spread with the fungus from northwest to southeast across Central America, beginning in the early 1980s. But this wave is still being evaluated. Federico Bolaños at the University of Costa Rica in San Pedro and his co-workers propose that it stems from biased sampling and from a failure to take into account all of the region’s amphibian die-offs. Moreover, this pattern does not rule out the importance of environmental change in these declines. Studies of such losses must test hypotheses that consider multiple factors.

The authors recognize the interplay of factors, but treat this as a complication that obscures a simple reality. If the chytrid plays a part in a die-off, they reason, then other forces are not needed. This is like attributing a car crash to excessive speed and deciding that other contributing issues, such as alcohol consumption, need not be considered. Such a deterministic view of causation, blind to probabilistic influences, has no place in ecology. Nor is the simplest explanation necessarily the best. Collins and Crump invoke Occam’s ‘law of parsimony’, but were Occam here today, he might quiz them about their assumptions.

For one, is the chytrid fungus really deadly?

**“Amphibians are telling us one thing: Earth’s life-support system is in trouble.”**



Amphibians such as this Tapichalaca tree frog are seen as bellwethers of environmental deterioration.

The resulting disease is sometimes lethal but often it is not, for reasons that go beyond genetics and history — environment matters in disease outcomes. Laboratory studies can be misleading: often they put amphibians under stress or erase the microclimatic heterogeneity that helps the animals to fend off disease. Such experiments can also exclude microbes that might otherwise keep the chytrid in check. Add to the picture how little we know about this fungus in the wild, especially outside its amphibian hosts, and the assumption that it is inherently lethal to many species becomes indefensible.

The chytrid’s impact may instead depend on environmental changes. Collins and Crump’s selection of published work and quoted opinions downplays such links. Nevertheless, studies show that increasingly extreme climate and weather, together with land-use change, pollution, ultraviolet (UV) radiation and species invasions, are degrading amphibian health in many regions. These factors can interact and their effects may cross the boundaries of protected areas. Associations between climatic changes and reduced amphibian survival have been found in western and eastern North America, Central and South America,

Australia, Spain, Italy and England.

Evidence showing that amphibians have undergone climatic stress before chytrid outbreaks have taken place challenges the idea of a lone-killer fungus, yet Collins and Crump fail to acknowledge this. Instead, they argue that such evidence casts doubt on the largely untested ‘chytrid-thermal-optimum’ hypothesis, which proposes that microhabitat temperatures in many places are shifting towards the chytrid’s optimum, thereby favouring this pathogen’s growth and hindering amphibian defences. However, the climatic-stress hypothesis is compatible with this model. And contrary to the authors’ claims, neither hypothesis assumes that the fungus is native to regions that are experiencing die-offs. Climate may influence a disease regardless of the pathogen’s place of origin, and the presence of exotic microbes makes global warming an even greater threat.

The authors’ narrow thinking biases their ‘road map’ for future research on amphibian declines. Consider UV radiation, which harms amphibians and interacts with pathogens, chemical pollution and climate in ways that scientists are only beginning to study. For example, global warming alters patterns of cloud cover, reducing UV exposure in some places and increasing it in others. By oversimplifying the

issue, the authors suggest that UV radiation can be removed from the list of potential causes of amphibian die-offs.

Collins and Crump offer little comparison between amphibian declines and losses affecting other species. Organisms from lodgepole pines to African lions are dying en masse because of disease shifts that are linked to global warming and other environmental problems. For instance, in the wake of the European heatwave of 2003, many millions of invertebrate sea creatures, including sponges, molluscs and corals, died along several thousand kilometres of the northwest Mediterranean coastline. After studying a soft coral known as the red gorgonian, marine biologists Marc Bally and Joaquim

Garrahou of the University of the Mediterranean in Marseille, France, reported that the high temperatures favoured certain bacteria that ravaged the coral. In Italy, a team led by Ines Di Rosa at the University of Perugia and Daniele Canestrelli at the University of Tuscia in Viterbo proposes that the heatwave had similar consequences for some of Italy's amphibians.

By asking how long it will be before amphibians are safe from the chytrid fungus, *Extinction in Our Times* misses the bigger picture. The interacting changes threaten many life forms. Di Rosa, Canestrelli and their colleagues are studying how climate change and pathogen movement may conspire to cause amphibian die-offs. In Costa Rica, where such losses

have accompanied climate-related changes in populations of birds, reptiles, mammals, insects and plants, we are studying orchids that have recently begun to die in the cloud forest. Amphibians belong to a chorus of canaries telling us one thing: Earth's life-support system is in trouble. ■

**J. Alan Pounds** is resident scientist at the Tropical Science Center's Monteverde Cloud Forest Preserve, Santa Elena, Puntarenas 5655-73, Costa Rica. **Karen L. Masters** is director of the Monteverde Program on Sustainability and the Environment of the Council for International Educational Exchange, Santa Elena, Puntarenas 5655-26, Costa Rica.  
e-mail: goldtoad@racsa.co.cr

## Newton and the money men

**Newton and the Counterfeiter: The Unknown Detective Career of the World's Greatest Scientist**

by Thomas Levenson

Houghton Mifflin Harcourt/Faber and Faber: 2009. 336 pp. \$25/£20

After the Glorious Revolution of 1688, in which King James II of England was overthrown by a union of Parliamentarians led by William of Orange, the English government found itself in dire financial straits. It had joined the War of the Grand Alliance against France in 1689, and was struggling to fund its army in a conflict that was to last for another 8 years. To make matters worse, the country was suffering from a lack of good coinage.

As Thomas Levenson explains in his engaging book *Newton and the Counterfeiter*, the government turned to an unlikely hero to save the nation from financial calamity — Isaac Newton.

In the 1660s, the English government had carried out a programme to machine-mill the edges of coins to dissuade unscrupulous operators from 'clipping' the edges and melting down the clippings for personal gain. But by the 1690s, many milled coins had dropped out of circulation, partly because their face value was less than the value of the silver they were made from. This caused a crisis for the Treasury, which would not take clipped coins as payment for tax because they contained less silver. So in 1696, the Treasury resolved to take £7-million-worth of non-milled coins out of circulation (today's circulation is £3.5 billion or US\$5.7 billion), melt them down and re-coin them with milled edges.

Newton was brought in to manage this operation. An unlikely appointee for the role, he was at that time enjoying fame as the author of *Principia Mathematica*, his seminal work on the foundations of physics, and had just embarked on a radical change of career as a politician. Newton became the Member of Parliament for the University of Cambridge in the Convention Parliament of 1689, formed in the wake of James II's departure. But his efforts to acquire a senior public position in London came to nothing until early 1696. Then, with the backing of his patron Charles Montagu — Chancellor of the Exchequer and 1st Earl of Halifax — Newton was awarded the position of Warden of the Royal Mint.

Although the job had been treated as a sinecure by most of his predecessors, Newton took it on with vigour. He masterfully oversaw the great re-coinage and, after overcoming

his initial revulsion, prosecuted with relish the clippers and 'coiners', or counterfeiters, who were partly responsible for the disarray of the country's currency. It wasn't long before his role brought him up against the arch-counterfeiter and forger, William Chaloner, whose skill and success in faking French pistoles (gold coins) and English guineas had quickly taken him from poverty to riches.

The book documents the entertaining relationship between these two geniuses and the different worlds they inhabited. Although their story is well known to historians of science, Levenson's account adds substantially to our knowledge of the social and political background against which it played out. The author manages to unpick many of the tangled and morally ambiguous webs that made up the metropolitan counterfeiting culture of that era, and shows — impressively, given the scant sources available — how Chaloner pulled off many of his brazen schemes.

It is an enthralling tale. At one point, Chaloner became wealthy enough to live in a large house in central London, but just as quickly lost whatever fortune he had made. He bounced back in his typical extraordinary fashion. In February 1697, he managed to convince a Parliamentary committee that was investigating alleged abuses at the Royal Mint — the allegations had come from Chaloner himself — that he could oversee a much more efficient way of producing coinage than the method that was in use.

Newton showed that Chaloner's scheme was unworkable. However, within a year the trickster had distributed a document making further accusations of corruption against members of the Mint, this time alluding to the activities of the warden himself. Again his claims were taken seriously, and Newton and others were



As Warden of the Royal Mint, Isaac Newton used his genius to investigate and convict a similarly intellectual counterfeiter.



investigated by the most senior members of the Treasury. Newton then put all of his energies into preparing a careful case against his rival, personally interrogating a number of Chaloner's former associates, and at a trial in March 1699 secured a conviction against him for counterfeiting. Chaloner feigned madness for a time — his final sham — but was hanged the same month.

*Newton and the Counterfeiter* contains the odd error of fact, and Levenson is on sticky historical ground with his claims about the wider contexts

of the events he describes. A more conspicuous drawback is the author's failure to consider Newton's feud with Chaloner in the light of his battles with other luminaries of that era, such as Robert Hooke and Gottfried Leibniz, or even Newton's relentless denunciation of the fourth-century Saint Athanasius — in his view, the propagator of the corrupt doctrine of the Holy Trinity. Privately, Newton expended a vast amount of time examining the morals and actions of people such as Athanasius, and finding them guilty of crimes against Christianity.

Newton may have known next to nothing about prosecuting clippers and coiners, as Levenson correctly notes, but once it became personal, Chaloner stood little chance against a man who spent much of his life cutting much larger intellectual heavyweights down to size. ■

**Robert Iliffe** is professor of intellectual history and the history of science in the Department of History, University of Sussex, Brighton BN1 9RH, UK; author of *Newton: A Very Short Introduction*; and editorial director of the Newton Project ([www.newtonproject.sussex.ac.uk](http://www.newtonproject.sussex.ac.uk)).

## Florence's observatory restored

### Torrino della Specola

Museum of Natural History, Florence

Opens 7 November

Newly restored, an elegant eighteenth-century astronomical observatory in Florence, Italy, reopens to the public this week after 135 years.

Its builder, Italian polymath Felice Fontana, had rushed from Pisa across the hills of Tuscany, bursting with zeal. The Grand Duke Pietro Leopardo had chosen him to direct his new Museum of Physics and Natural History in Florence, which opened to the public in 1775. Fontana now had carte blanche to realize his Utopian dream: a museum in which all that was known of nature could be brought into a single building for the edification of the people, and for scientists to make yet more discoveries. It would be a pinnacle of Enlightenment endeavour, and Fontana intended it to be on a par with Florence's unequalled fine-art collections.

The Grand Duke commissioned the restructuring of the medieval Torrigiani palace for the museum, ensuring its prominent position between two Renaissance landmarks: the monumental Pitti palace and the formal, sculpture-filled Boboli gardens.

Fontana threw himself into acquiring collections based on natural history, botany, mineralogy and more — as well as commissioning the wax anatomical and botanical models for which Florence is justly famous (see *Nature* 452, 414; 2008). Astronomy, neglected by the city since Galileo's death there in 1642, was to be a major activity. So the Torrigiani palace was reinforced to bear the weight of a 35-metre-high observation tower, the Torrino della Specola.

This was an unpopular move. Even the observatory's first director, Domenico de Vecchi, was outspokenly critical of its logistical virtues. "For all its elegance, it is not the most favourable for observations, nor the most comfortable for

observers," he wrote in 1808, a year after he was appointed director.

Indeed, few astronomers approved. The modern trend was to build observatories away from towns and above the mist, on top of hills where there was plenty of space. The observatory tower was simply cramped — the architect had built it without consulting astronomers and had failed to provide space for them to



Stuccoed storks flank one of the Florence observatory's rooms.

retire to their books or writing tables. Annoyingly, the perfect hill was there for the taking — Arcetri, overlooking Florence, where Galileo had made his last home. Galileo's residence, known as Il Gioiello, or The Jewel, was restored and reopened earlier this year.

But Fontana was unshakably wedded to his vision of the all-encompassing museum and science centre. He resisted incessant calls to relocate the observatory. Fontana died in 1805, but more than 60 years passed before pragmatism finally won out and in 1872, the astronomers moved to a new observatory on Arcetri, which still operates today. By then, Fontana's Utopia was being broken up, with many collections being distributed to other sites in Florence, including the university. Only the natural-history and anatomical collections remained in the museum.

The observatory tower remained closed, but has now been renovated to its original glory. It reopens as an extension to the museum on 7 November, in this International Year of Astronomy that marks the 400th anniversary of Galileo's first observations with a telescope.

There are three good reasons to visit. First is the exhibition. Paying homage to Fontana's vision of the all-inclusive scientific centre, the tower displays representative artefacts from the original collections — ornaments acquired by Captain James Cook; two bewilderingly life-like magnolia and lotus flowers made of wax; a Medici collection of worked gemstones; an ancient herbarium; some fossils, a pair of taxidermically prepared lion monkeys; a few historic telescopes and a couple of still-life paintings by Bartolomeo Bimbi. The displays are not extensive — the space hasn't got any larger in the past two centuries — but they are fine examples.

Second, the architecture is a work of art — octagonal rooms, gracious windows and elegant, narrow spiral staircases. Most astounding is the Meridian Room, whose supporting arches are stuccoed with slender storks in the act of taking flight. A marble meridian inlaid into the floor, decorated in copper and silver, indicates the hour and time of year. When astronomers used this room for observations, a narrow slit in the walls and roof could be opened to expose a 180° slice of the sky. Eight large windows and an outside terrace allowed an unobstructed view of the whole sky.

The third reason to visit the tower is its superb location. Its elevation allows a unique and commanding view of one of the world's most beautiful cities — a panorama that has not been seen by the public for 135 years. ■

**Alison Abbott** is *Nature's* senior European correspondent.

For more details on the observatory tower, see [www.msn.unifi.it](http://www.msn.unifi.it). See [go.nature.com/V52qT5](http://go.nature.com/V52qT5) for articles in the hidden treasures series.

## IMMUNOLOGY

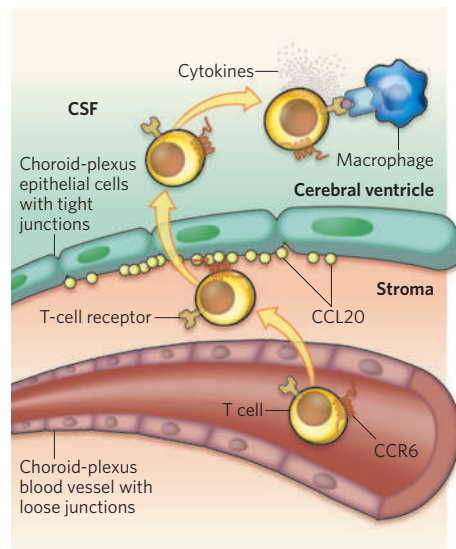
## In the beginning

Richard M. Ransohoff

**Immune cells cross the inflamed blood–brain barrier. But it's unclear how brain inflammation begins before immune-cell entry. Studies of a model of multiple sclerosis start to solve this 'chicken and egg' conundrum.**

Multiple sclerosis, a disorder of the central nervous system (CNS), is characterized by damage to myelin, the insulating membrane that surrounds nerve fibres. The lesions found in multiple sclerosis are inflamed, but don't contain an evident pathogen that might have triggered an immune reaction; hence, it has been proposed that the disorder is caused by an autoimmune process, in which the body's defence system turns against its own tissue. To initiate inflammation and damage, T lymphocytes must enter the tissue (parenchyma) of the CNS<sup>1,2</sup> by crossing the blood–brain barrier, which is comprised of usually impenetrable tight junctions between the endothelial cells that line the cerebral blood vessels. But T cells bind to vessels of the blood–brain barrier only when inflammation activates the blood-vessel endothelium<sup>3</sup>, so how and where immune cells first enter the uninfamed CNS had posed a mystery. Two studies, one by Reboldi *et al.*<sup>4</sup> in *Nature Immunology*, and the other by Bartholomäus and colleagues<sup>5</sup> on page 94 of this issue, vividly illuminate the elaborate processes by which autoimmune reactions begin in the CNS.

Studies of experimental autoimmune encephalomyelitis (EAE), an animal model of multiple sclerosis, have been invaluable in shedding light on the mechanisms underlying CNS autoimmunity. EAE is generated by subcutaneous injection of fragments of myelin proteins — the antigens against which T cells are believed to react in multiple sclerosis. About two weeks after immunization, immune white blood cells, such as lymphocytes and monocytes, accumulate in the animal's spinal cord, and motor weakness develops. One implication of the EAE model is that some of the events that result in EAE-related tissue inflammation might also be involved in CNS immune surveillance — the process by which T cells 'patrol' the CNS to identify and destroy pathogens. In both cases, T cells educated in the periphery by exposure to antigen seem to identify similar, or identical, antigens in the brain<sup>6,7</sup>. But whereas in immune surveillance the outcome is protective, in EAE, T cells attack self tissue, causing disease. In either case, though, the mechanisms by



**Figure 1 | The first wave of immune-cell entry.** Reboldi and colleagues<sup>4</sup> investigate initial T-cell entry into the uninfamed CNS in mice developing an animal model of multiple sclerosis. CCR6-expressing effector and memory T cells escape from choroid-plexus blood vessels into the supporting tissue stroma, then migrate towards choroid-plexus epithelial cells, which express the chemoattractant molecule CCL20. Signalling by CCL20 through CCR6 allows T cells to cross the tight junctions between the choroid-plexus epithelial cells (the blood–CSF barrier) and enter the cerebral ventricles, from which they migrate to the subarachnoid space. Restimulation of T cells by macrophages in the subarachnoid space leads to the production of growth and inflammatory factors (cytokines), increased T-cell numbers and stimulation of the blood–brain vasculature, which allows the second wave of inflammatory-cell CNS invasion.

which T cells first enter the CNS have not been clear.

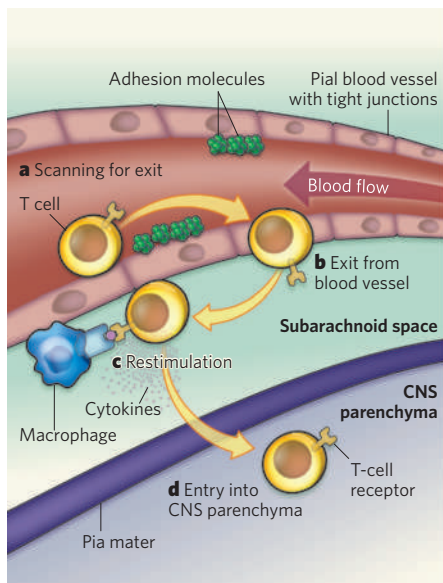
Reboldi *et al.*<sup>4</sup> investigate this first stage of immune-cell entry by focusing on a subset of T cells, T<sub>H</sub>17 cells, that are thought to be involved in autoimmune inflammation. On their surface, T<sub>H</sub>17 cells express CCR6, a receptor for CCL20 — one of a family of molecules that guide the migration of white blood

cells into tissues. The authors find that mice lacking CCR6 (CCR6-knockout mice) are resistant to developing EAE: such mice do not develop CNS-tissue inflammation or motor weakness after immunization with myelin proteins, despite showing normal T<sub>H</sub>17-cell responses. Furthermore, video microscopy shows that T<sub>H</sub>17 cells from CCR6-knockout mice behave like those of normal mice in that they bind to inflamed, but not to uninfamed, vessels of the blood–brain barrier. Unexpectedly, transfer of small numbers of T cells from unimmunized normal mice into CCR6-knockout mice before immunization restores susceptibility to EAE — the transferred CCR6-expressing T cells are sufficient to initiate inflammation, after which CCR6-deficient T cells easily invade the now-inflamed CNS.

How can these intriguing findings be explained? Careful examination of CNS tissues by Reboldi *et al.*<sup>4</sup> provides some answers. After CCR6-knockout mice have been immunized with myelin proteins, CCR6-deficient T cells accumulate in the choroid plexus, a region of the brain that is rich in capillary blood vessels and that secretes cerebrospinal fluid (CSF) into the cerebral ventricles. CSF circulates from the ventricles through the subarachnoid space between the outer arachnoid and inner pial membranes, bathing and buffering the CNS tissues. CSF-secreting choroid-plexus epithelial cells express CCL20, and the authors show that CCL20 must signal through CCR6 for T<sub>H</sub>17 cells to enter the CSF through the choroid plexus (Fig. 1). The entry of T<sub>H</sub>17 cells, in turn, begins the process that culminates in brain and spinal-cord inflammation.

The authors' results<sup>4</sup> agree with findings indicating that, before the onset of EAE, memory T cells (those previously exposed to antigen) are reactivated by antigen in the subarachnoid space<sup>8</sup>. Laudably, Reboldi and colleagues<sup>4</sup> also studied human brain tissue and find that CCL20 decorates choroid-plexus epithelial cells in humans as well as mice. Furthermore, CSF samples from healthy humans are replete with memory T cells<sup>9,10</sup>. Taken together, these data<sup>4,8–10</sup> establish that CNS immune surveillance is carried out by memory T cells





**Figure 2 | The second wave.** Bartholomäus and colleagues<sup>5</sup> use two-photon imaging to visualize the accumulation of T cells in the CNS after activation of the blood–brain barrier. T cells roll along, and then adhere to, the inner wall of activated cerebral blood vessels (expressing adhesion molecules) that lie next to the pia mater — the inner membrane lining the surface of the brain and spinal cord. **a**, T cells crawl along the vessel wall, mainly against the direction of blood flow, presumably in search of a suitable exit site. **b**, After crossing into the subarachnoid space, these cells move along the outer wall of the blood vessel, where they come into contact with local resident macrophages, which present them with specific antigen. **c**, Antigen recognition results in T-cell restimulation and production of abundant cytokines. **d**, T cells then detach from the blood-vessel surface and enter the CNS parenchyma, where they cause tissue damage.

patrolling the subarachnoid space, where crucial restimulation of T cells takes place<sup>10,11</sup> after their initial encounter with antigen in the periphery. In the case of EAE, Reboldi and colleagues<sup>4</sup> conclude that inflammatory cells invade the CNS in two waves. The first wave is through the choroid plexus, leading to T-cell restimulation by antigen in the subarachnoid space, early CNS inflammation and stimulation of the blood–brain-barrier vasculature (Fig. 1)<sup>12</sup>. Activation of the blood–brain barrier then initiates the second wave: a large-scale inflammatory-cell invasion of the CNS parenchyma, with resulting motor weakness<sup>4,13</sup>.

Bartholomäus *et al.*<sup>5</sup> provide a dramatic account of the initial moments of this second wave of CNS inflammatory-cell entry. They study an adoptive-transfer model of EAE in which rats are injected with activated anti-myelin T cells, developing signs of the disease three to four days later. Intravenous injection of activated T cells stimulates the cerebral blood vessels, enabling T cells to bind<sup>14</sup>, and obviating the need for the first wave of T-cell entry

across the choroid plexus. Bartholomäus *et al.* use sophisticated two-photon imaging to visualize labelled T cells as they move along and exit the blood–brain barrier's pial vessels on the surface of the spinal cord.

The authors witness a fantastic exhibition of lymphocyte gymnastics (see the movies in Supplementary Information to ref. 5). Disease-causing T cells first adhere to the inner walls of the pial vessels and then crawl in continuous contact with activated endothelial cells, most often in the opposite direction to the blood flow. Crawling allows these cells to identify vascular exit sites<sup>15</sup> (Fig. 2a). After crossing the blood-vessel wall (Fig. 2b), the lymphocytes move along the outer surface of the vessel, encountering an array of antigens displayed by antigen-presenting cells, including macrophages. If the T cell recognizes a specific antigen, it remains in prolonged contact with the antigen-presenting cell and produces abundant growth or inflammatory factors (cytokines) (Fig. 2c). Last, the cells detach from the outer surface of the blood vessel and enter the spinal cord, travelling most often alongside penetrating vessels (Fig. 2d). In the spinal cord, they initiate tissue injury. Bartholomäus *et al.* additionally show that activated T cells that are reactive to non-brain antigens also accumulate in, and crawl along, the inner walls of spinal-cord vessels. This finding shows that reaction against CNS antigens is not required for intravascular crawling, which seemed relatively selective for spinal-cord vessels in the authors'

studies. However, extravascular restimulation with antigen is required for lymphocyte entry into the CNS parenchyma.

Bartholomäus and colleagues' work<sup>5</sup> is replete with such interesting concepts and findings that its careful perusal will undoubtedly yield many more insights. This coherent molecular close-up of the first steps in CNS autoimmunity brings with it the promise of new therapeutic strategies.

Richard M. Ransohoff is in the Neuroinflammation Research Center, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio 44195, USA. e-mail: ransohr@ccf.org

1. Glabinski, A. R., Tani, M., Tuohy, V. K., Tuthill, R. J. & Ransohoff, R. M. *Brain Behav. Immun.* **9**, 315–330 (1995).
2. Gerard, C. & Rollins, B. J. *Nature Immunol.* **2**, 108–115 (2001).
3. Piccio, L. *et al. J. Immunol.* **168**, 1940–1949 (2002).
4. Reboldi, A. *et al. Nature Immunol.* **10**, 514–523 (2009).
5. Bartholomäus, I. *et al. Nature* **462**, 94–98 (2009).
6. Galea, I., Bechmann, I. & Perry, V. H. *Trends Immunol.* **28**, 12–18 (2007).
7. Ransohoff, R. M. *Nature* **457**, 155–156 (2009).
8. Kivisakk, P. *et al. Ann. Neurol.* **65**, 457–469 (2009).
9. Kivisakk, P. *et al. Proc. Natl Acad. Sci. USA* **100**, 8389–8394 (2003).
10. Ransohoff, R. M., Kivisakk, P. & Kidd, G. *Nature Rev. Immunol.* **3**, 569–581 (2003).
11. Kawakami, N. *et al. J. Exp. Med.* **199**, 185–197 (2004).
12. Brown, D. A. & Sawchenko, P. E. *J. Comp. Neurol.* **502**, 236–260 (2007).
13. Axtell, R. C. & Steinman, L. *Nature Immunol.* **10**, 453–455 (2009).
14. Xu, H. *et al. J. Neuroimmunol.* **142**, 47–57 (2003).
15. Schenkel, A. R., Mamdough, Z. & Muller, W. A. *Nature Immunol.* **5**, 393–400 (2004).

## EARTH SCIENCE

# Lasting earthquake legacy

Tom Parsons

**Earthquakes occur within continental tectonic plates as well as at plate boundaries. Do clusters of such mid-plate events constitute zones of continuing hazard, or are they aftershocks of long-past earthquakes?**

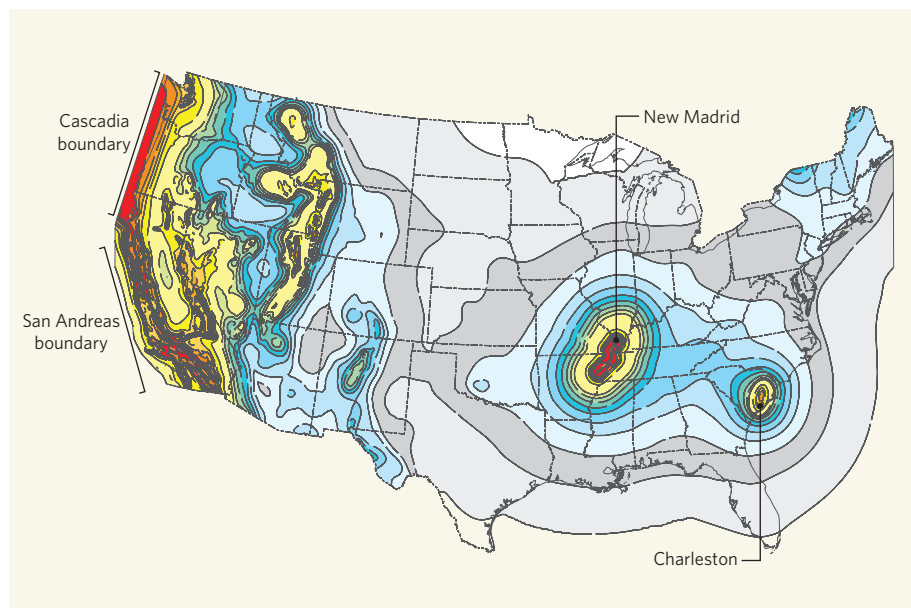
Early on the morning of 16 December 1811, an earthquake of about magnitude 7 shook the centre of the United States around a small town on the Mississippi called New Madrid. By 7 February 1812, it had triggered three more shocks of similar magnitude<sup>1</sup>. The earthquakes broke a set of faults along the Arkansas, Missouri and Tennessee state boundaries, apparently reactivating an ancient rift in the interior of a continental tectonic plate<sup>2</sup>.

On 31 August 1886, a magnitude-7 shock struck Charleston, South Carolina; low-level activity continues there today. One view of seismic hazard is that large earthquakes will return to New Madrid and Charleston at intervals of about 500 years<sup>3</sup>. With expected ground motions that would be stronger than average<sup>4</sup>, that prospect produces estimates of earthquake hazard that rival those at the plate

boundaries marked by the San Andreas fault and Cascadia subduction zone<sup>3,4</sup>. The result is two large 'bull's-eyes' on the US National Seismic Hazard Maps (Fig. 1) — which, for example, influence regional building codes and perceptions of public safety.

But what if earthquakes are not always going to return to mid-continental locations such as New Madrid or Charleston? From data on activity at New Madrid and elsewhere in the world, Stein and Liu (page 87 of this issue)<sup>5</sup> identify a global inverse correlation between aftershock duration and tectonic deformation rates. They argue that localized high rates of mid-continental activity, taken to reflect steady-state plate deformation, may instead be very long-lived aftershock sequences.

As first noted by Omori<sup>6</sup> in 1894, and now known as Omori's law, aftershock rates



**Figure 1 | Probabilistic assessment of seismic hazard in the United States<sup>3</sup>.** Warm colours show regions with the highest probability of strong shaking, which tend to be located near, or influenced by, major plate boundaries. Earthquake clustering in mid-continental areas such as New Madrid and Charleston is here interpreted to reflect continuing deformation sufficient to produce frequent large earthquakes. This expectation, combined with the anticipated ground response, produces hazard estimates that rival those for plate-boundary zones on the west coast. Stein and Liu<sup>5</sup>, however, suggest that low-level, mid-continental seismic activity can be attributed to a long-lived sequence of aftershocks.

are highest immediately after a mainshock, and then decay as a function of inverse time (Fig. 2). Stein and Liu<sup>5</sup> demonstrate that aftershocks decay very slowly where there is little tectonic deformation, and can persist for hundreds of years. By contrast, aftershock rates at fast-deforming plate boundaries return to the background level within about ten years.

Rate-and-state friction theory holds that fault failure depends on competition between initially very slow slip speed (rate) and the evolving duration (state) of a population of

contact points in a fault. Stein and Liu's observations confirm a major prediction of that theory — that the aftershock decay period on a fault should depend directly on the rate at which tectonic stress accumulates<sup>7</sup>. Declining aftershock sequences at New Madrid and Charleston, along with little steady-state deformation, would mean that earthquake hazard rates are declining as well.

We have independent methods of assessing whether plate deformation is continuing. On the one hand, palaeoseismology — a kind

of forensic seismology that reads geological clues left by prehistoric earthquakes — has been used to interpret quake clusters that occurred around AD 900 and AD 1450 at New Madrid<sup>8</sup>. When combined with the 1811–12 series, the upshot of this analysis was the estimate of a 500-year recurrence interval used for hazard mapping<sup>3</sup>. On the other hand, space geodesy — the use of a space-based reference frame to monitor current crustal movements — reveals present-day deformation rates in the New Madrid zone that are indistinguishable from those of the rest of the continental interior, and that have been interpreted as New Madrid shutting down<sup>9</sup>. Digging deeper into the

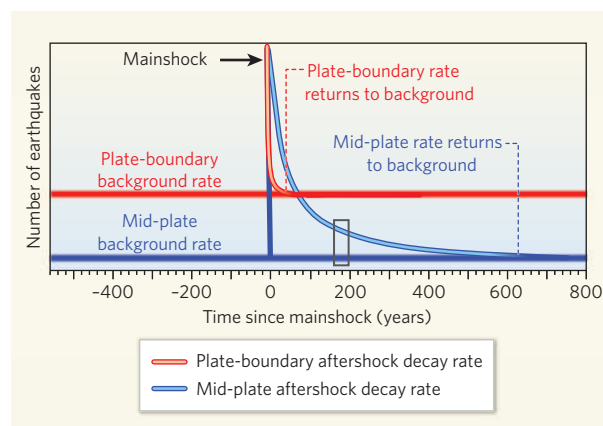
palaeoseismic record indicates temporal earthquake clusters that are followed by millennia of quiescence<sup>10</sup>, a finding that reconciles these contrasting observations if we are now entering a quiet period.

This is where geological time versus human time comes in. Meteorologists have the benefit of having watched hundreds of seasonal cycles to support their weather forecasts. Earthquake forecasters, however, have not yet witnessed a complete seismic cycle in slowly deforming continental interiors anywhere on Earth. The task is akin to predicting a full-year's weather based on watching one week in January. Stein and Liu's study<sup>5</sup> inevitably suffers from this problem; it is difficult to accurately calculate a long aftershock decay trend from a brief observation time. A limited view of earthquake history can actually be dangerous: despite the huge shocks that recently caught many by surprise in Sumatra and in Wenchuan, China, there is still a tendency to place undue emphasis on known past events in hazard assessments when trying to see the future.

A better understanding of mid-continental seismicity, one that avoids excessive influence from events in the recent past, will require that historically unbroken crustal weak points are not overlooked. Fault mapping and palaeoseismology at unconventional locations may be required to find out if places such as New Madrid and Charleston are special. At the same time, we must discriminate between the signals of steady-state deformation and the slowly fading legacy of historical earthquakes.

The large uncertainties arising from our brief glimpse into mid-continental earthquake cycles are unavoidable. Forecasters will thus have to follow multiple possibilities, including those with Omori-law time-decaying hazard estimates, through a logic-tree structure. With that approach, there is the prospect of arriving at a practical solution that properly balances the available resources for earthquake mitigation and the distribution of earthquake-resistant building structures in mid-continental regions.

Tom Parsons is at the US Geological Survey, 345 Middlefield Road, Menlo Park, California 94025, USA.  
e-mail: tparsons@usgs.gov



**Figure 2 | Comparison of aftershock decay rates from plate-boundary and mid-plate regions.** Mainshocks cause temporarily heightened earthquake rates of a duration controlled by the rates of tectonic deformation<sup>5</sup>. The black box shows the modern instrumental observation window on a long-lived decay sequence of mid-plate aftershocks. That window has been open for only about 40 years, illustrating the difficulties in distinguishing aftershocks from the lower, steady-state background rate of earthquake occurrence.

1. Mueller, K., Hough, S. E. & Bilham, R. *Nature* **429**, 284–288 (2004).
2. Braille, L. W., Keller, G. R., Hinz, W. J. & Lidiak, E. G. *Tectonics* **1**, 225–237 (1982).
3. Frankel, A. et al. *US Geol. Surv. Open-File Rep.* 02-420 (2002).
4. Frankel, A. *Seismol. Res. Lett.* **75**, 575–586 (2004).
5. Stein, S. & Liu, M. *Nature* **462**, 87–89 (2009).
6. Omori, F. *Rep. Imperial Earthquake Investigation Committee* **2**, 103–109 (1894).
7. Dieterich, J. H. *J. Geophys. Res.* **99**, 2601–2618 (1994).
8. Tuttle, M. P. et al. *Bull. Seismol. Soc. Am.* **92**, 2080–2089 (2002).
9. Calais, E. & Stein, S. *Science* **323**, 1442 (2009).
10. Holbrook, J., Autin, W. J., Rittenour, T. M., Marshak, S. & Goble, R. J. *Tectonophysics* **420**, 431–454 (2006).



## CANCER

# A tumour gene's fatal flaws

Julian Downward

**Mutations in RAS genes are common in human tumours, but RAS has proved impossible to target with drugs. Its associated NF- $\kappa$ B signalling pathway, however, may turn out to be this tumour gene's Achilles heel.**

RAS is one of the most commonly mutated gene families in human cancers — one of its three members (*KRAS*, *HRAS* and *NRAS*) is mutated in about 20% of human tumours. Attempts to target mutant RAS proteins directly with small-molecule inhibitors have so far proved unsuccessful, so there has been considerable interest in finding signalling pathways that function downstream of RAS and whose blockade might be selectively toxic to tumour cells. Two papers in this issue provide evidence that targeting one such pathway, the NF- $\kappa$ B signalling pathway, may be an effective approach to treat RAS-mutant tumours such as lung cancers. Barbie *et al.*<sup>1</sup> (page 108) identify a component of the NF- $\kappa$ B pathway as a potential target in RAS-mutant cancer cells, and Meylan *et al.*<sup>2</sup> (page 104) show that inhibition of NF- $\kappa$ B signalling impairs tumour formation in a mouse model of RAS-induced lung cancer.

The use of the RNA-interference technique to selectively inhibit gene expression, together with knowledge of the full sequence of the human genome, has made possible large-scale functional genomic screens. In these, each gene in the genome (or at least a significant proportion of genes) is silenced one by one, and the effect on cell function is assayed. This approach has recently been used<sup>3,4</sup> to investigate which genes, when silenced, kill cells bearing mutant RAS but not cells that lack this mutation — so-called synthetic lethal interactions. Barbie *et al.*<sup>1</sup> looked for synthetic lethal interactions in a panel of cell lines, some of which had activating mutations in *KRAS*. The authors inhibited genes thought to be important for the development of cancer and identified several genes whose reduced activity seemed to selectively kill the RAS-mutant cells. After *KRAS* itself, silencing the gene *TBK1*, an upstream regulator of the NF- $\kappa$ B pathway, was most effective for selectively killing RAS-mutant cells. *TBK1* is thought to activate NF- $\kappa$ B by phosphorylating I $\kappa$ B, an NF- $\kappa$ B inhibitor. It also activates other transcription factors involved in inflammatory responses, such as IRF3 and IRF7 (ref. 5).

The NF- $\kappa$ B transcriptional program controls a multitude of cell functions, most notably the regulation of cell death and inflammation<sup>6</sup>. This is by no means the first time that this pathway has been linked to RAS-mediated tumour formation. NF- $\kappa$ B signalling has been shown<sup>7</sup> to suppress the tendency of mutant RAS to cause stress-induced cell death, in part by

undermining cell-cycle checkpoints. Indeed, *TBK1* has been shown<sup>8</sup> to have a crucial role downstream of RAS, being activated by RalB, a small GTP-binding protein that is controlled by RalGDS, one of the downstream effectors of RAS (Fig. 1).

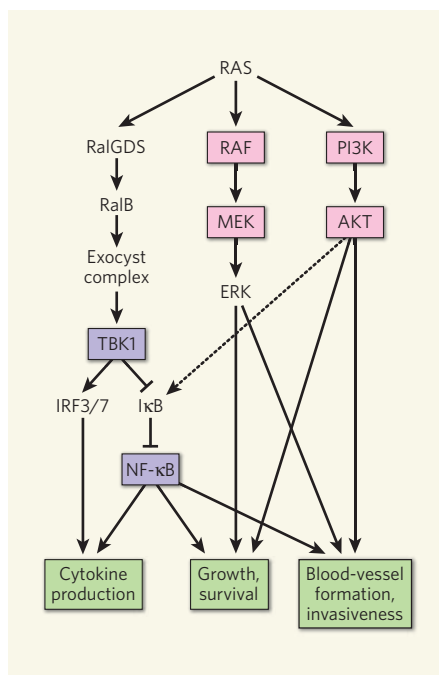
The finding that RAS-mutant cells are dependent on *TBK1* *in vitro* indicates that it would make sense to attempt to inhibit the NF- $\kappa$ B pathway in these tumour types. However, the links between NF- $\kappa$ B and the inflamma-

tory response, together with our growing awareness of the importance of inflammation in the pathogenesis of cancer<sup>6</sup>, suggest that the role of NF- $\kappa$ B in cancer is probably poorly modelled by tissue-culture experiments using homogeneous cell populations. What has been lacking so far are data from an animal model of cancer that is characterized by RAS mutation. These data have now been provided by Meylan and colleagues<sup>2</sup>. The authors engineered mice in which the mutated cancer-promoting gene (oncogene) *KRAS* is expressed at the same time as the *p53* tumour-suppressor gene is deleted in the lung; the result is the development of an aggressive lung cancer. Simultaneous expression of an inhibitor of the NF- $\kappa$ B pathway, the I $\kappa$ B 'super repressor', causes a massive reduction in the number and size of lung tumours formed. Also, inhibition of NF- $\kappa$ B in established lung tumours using an inducible form of the I $\kappa$ B super repressor slowed their growth, although tumours did not regress.

The results of this first foray into the targeting of NF- $\kappa$ B in *KRAS*-mutant lung tumours are encouraging, but much remains to be done before we can assess the likelihood of the success of this approach in the clinic. The failure to cause regression of existing tumours compares unfavourably with studies using other inhibitors; for instance, a combination of small-molecule inhibitors targeting two major RAS-effector pathways, MEK and PI3-kinase (Fig. 1), caused regression of similar *KRAS*-induced lung tumours<sup>9</sup>. Small-molecule inhibitors have been developed against the NF- $\kappa$ B pathway — it will be interesting to see how these perform in Meylan and colleagues' mouse model<sup>2</sup>.

A consistent worry with targeting NF- $\kappa$ B in cancer has been that prolonged inhibition of the pathway will cause immunosuppression, so it may be necessary to limit the length and frequency of systemic treatments with such inhibitors. This was not an issue for Meylan and colleagues<sup>2</sup>, as the I $\kappa$ B super repressor is expressed only in tumour cells. Another concern is the harmful effects reported when NF- $\kappa$ B is suppressed in some tumour types, including one mouse model of human skin cancer driven by an activated RAS oncogene, in which inhibition of NF- $\kappa$ B led to tumour progression, rather than its regression<sup>10</sup>.

An issue worth considering is whether targeting *TBK1* might have unique advantages over other ways of inhibiting NF- $\kappa$ B signalling in RAS-mutant tumours. This question is difficult to answer right now, as the pathway displays considerable complexity and the function of some of the lesser-known components, such as *TBK1*, is still relatively poorly understood. However, the work of Barbie *et al.*<sup>1</sup> and Meylan *et al.*<sup>2</sup> clearly establish the importance of NF- $\kappa$ B signalling in RAS-induced tumour formation, and highlight the potential value of pharmacological targeting of this pathway in lung and possibly other common cancers. ■



**Figure 1 | The RAS and NF- $\kappa$ B signalling pathways.** Activation of mutated RAS in tumours stimulates many downstream signalling pathways, including the RAF–MEK–ERK kinase cascade, the phosphatidylinositol 3-kinase (PI3K)–AKT pathway and RalGDS proteins, which are nucleotide-exchange factors that activate small GTPases, such as RalB. RalB functions through the exocyst protein complex to stimulate *TBK1*, which activates NF- $\kappa$ B by phosphorylating its inhibitor, I $\kappa$ B. *TBK1* also activates other transcription factors, such as IRF3 and IRF7, which leads to the release of growth and inflammatory factors (cytokines). Barbie *et al.*<sup>1</sup> show that RAS-mutant cancer cells are dependent on *TBK1* expression, and Meylan *et al.*<sup>2</sup> demonstrate that inhibition of NF- $\kappa$ B inhibits RAS-induced lung-tumour formation in mice. Potential (purple) or actual (pink) drug targets for treatment of RAS-mutant tumours are indicated.

Julian Downward is at Cancer Research UK London Research Institute, 44 Lincoln's Inn Fields, London WC2A 3PX, UK.  
e-mail: downward@cancer.org.uk

1. Barbie, D. A. *et al.* *Nature* **462**, 108–112 (2009).
2. Meylan, E. *et al.* *Nature* **462**, 104–107 (2009).
3. Luo, J. *et al.* *Cell* **137**, 835–848 (2009).

4. Scholl, C. *et al.* *Cell* **137**, 821–834 (2009).
5. Hacker, H. & Karin, M. *Sci. STKE* 2006, re13 (2006).
6. Baud, V. & Karin, M. *Nature Rev. Drug Discov.* **8**, 33–40 (2009).
7. Mayo, M. W. *et al.* *Science* **278**, 1812–1815 (1997).
8. Chien, Y. *et al.* *Cell* **127**, 157–170 (2006).
9. Engelman, J. A. *et al.* *Nature Med.* **14**, 1351–1356 (2008).
10. Dajee, M. *et al.* *Nature* **421**, 639–643 (2003).

## MATERIALS SCIENCE

# Soft is strong

C. Austen Angell and Kazuhide Ueno

**The mechanisms that govern the rate at which glasses soften on heating have long been a mystery. The finding that colloids can mimic the full range of glass-softening behaviours offers a fresh take on the problem.**

For hundreds of years, scientists have wondered how glass-forming liquids solidify without crystallizing as they cool — or indeed without undergoing much change in structure at all. Glass artists take advantage of this continuous rigidification process to create and preserve liquid 'form' and 'action', fascinating onlookers as they turn glowing globs of hot molten glass into elegant shapes. The artists take care to use a glass-former that converts to a solid over a wide range of temperatures, so that they have time to work their magic. Other glass compositions make the transition from fluid to solid much more suddenly, and it would be impossible to work with them.

Glass-forming liquids that soften quickly on heating are described as being fragile, whereas those that don't are 'strong'. The same distinction is made to describe the behaviour of liquids under compression — fragile liquids 'jam up' suddenly into glasses under pressure, and strong ones don't. But we don't know why. On page 83 of this issue, Mattsson *et al.*<sup>1</sup> report that the same pattern of strength and fragility holds for certain colloidal suspensions, and that they do know why. This offers hope that colloids could be used as model systems to help describe liquid behaviour.

To explain Mattsson and colleagues' findings, we first need to remind ourselves of the difference between colloids and liquids. Liquids are composed only of molecules (or atoms or ions), which ceaselessly collide, oscillate and diffuse. The fissures and gaps that form transiently in liquids are empty of matter. Colloidal suspensions, on the other hand, consist of microscopic chunks of matter, usually — but not necessarily — spherical in shape, suspended in a liquid medium. The suspended matter would settle slowly to the bottom of the colloid if it were not for the impacts it makes with the endlessly oscillating molecules of the liquid. Instead, the particles undergo Brownian motion, diffusing randomly, much like molecules, and providing a source of wonder for generations of science students peering down microscopes.

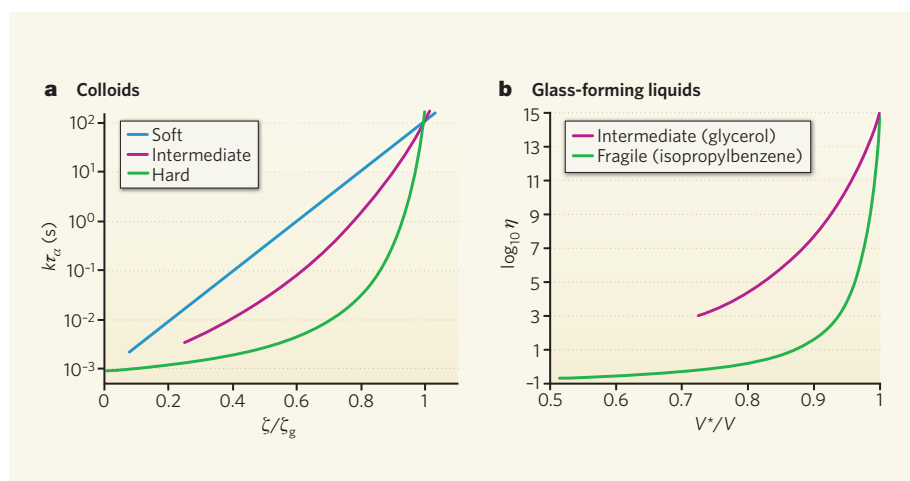
To describe all of this, one can measure the viscosity of the colloid, the diffusivity of the particles (a measurement of the particles' ability to diffuse through a particular medium) and the corresponding relaxation process that occurs when shear stress in the fluid is dissipated by viscous flow of liquid molecules or colloidal particles.

In most colloids that have been studied, the suspended particles are as solid as rock. As the concentration of the particles is increased (and so too the percentage of the colloid's volume that is occupied by particles, known as the packing fraction), the relaxation times of the colloids increase dramatically. The relaxation time reaches 100 seconds — the relaxation time that defines the glass transition point of a material — at a packing fraction of 0.58.

This value is called the glass transition packing fraction by analogy with the glass transition temperature. The behaviour of such colloids is remarkably similar to that of 'hard-sphere fluids' (the classical model of liquids in which the liquid particles are thought of as non-compressible spheres), for which the maximum packing fraction seems to be 0.64–0.65 (refs 2–5). The behaviour of hard-particle colloids turns out to be the fragile extreme of colloidal behaviour.

To get the equivalent of strong behaviour, Mattsson *et al.*<sup>1</sup> had to make their colloid particles soft. Such particles deform as they are packed closer together, tending to become cuboidal instead of spherical — a process brought about by a redistribution of the solvent inside the particles. To study the effects of different degrees of softness, the authors made three different kinds of particle from polymer networks that had different cross-linking densities. They were thus able to observe the effect of particle softness on the relaxation times of the colloids. The authors found that the variation of colloidal relaxation times with respect to particle concentration changes distinctively as the particles become harder (Fig. 1a).

So how do these data<sup>1</sup> connect with the fragility of glass-forming liquids? To find out we must go back to the 1990s, when King and colleagues<sup>6,7</sup> studied the effect of volume compression on the viscosity of certain non-crystallizing liquids. In a series of heroic experiments, they suspended balls of nickel in tiny volumes of liquid inside high-pressure diamond-anvil cells, then whirled the cells around in circles so that huge centrifugal forces acted on the balls. By measuring the distance rolled by the balls in a known period of time, King and



**Figure 1 | Colloidal relaxation times are analogous to liquid fragility.** **a**, Mattsson *et al.*<sup>1</sup> report that the variation of colloidal relaxation time,  $\tau_\alpha$ , with respect to colloidal particle concentration,  $\zeta$ , (scaled by the particle concentration at glass transition,  $\zeta_g$ ), depends on the softness of the particles.  $k$  is a scaling factor. **b**, The variation of viscosity,  $\eta$ , with volume,  $V$  (scaled by the volume  $V^*$  at glass transition) for certain glass-forming liquids was previously reported by King and colleagues<sup>7</sup>, some of whose data are reproduced here. The trends observed in **a** are strikingly similar to those of the viscosity–volume relationships:  $\tau_\alpha$  of the 'hard' colloids varies like the viscosity of fragile liquids (such as isopropylbenzene, shown);  $\tau_\alpha$  of colloids of intermediate softness varies like the viscosity of intermediate liquids (which have viscosities somewhere between fragile and strong, such as glycerol, shown). No comparable data for the viscosities of strong liquids are available. Viscosities were reported in centipoise.



colleagues calculated the viscosities of many liquids, over enormous ranges of viscosity and pressure, all with high precision (Fig. 1b). The variation in colloidal relaxation times measured by Mattsson *et al.*<sup>1</sup> is strikingly similar to that of King and colleagues' viscosity data<sup>6,7</sup>. In particular, the trends observed by Mattsson *et al.* in the variation of relaxation time with particle concentration as colloidal particles become harder are the same as the trends in the viscosity–volume relationships of increasingly fragile liquids (Fig. 1).

So, hard colloids are analogous to fragile liquids, whereas soft colloids can be compared to strong liquids. But why is soft strong? Mattsson *et al.*<sup>1</sup> offer the attractive explanation that

the elastic energy of colloidal particles — the energy stored when the particles are distorted — determines the fragility of colloids. This meshes nicely with other elasticity-based theories<sup>8</sup> that have been proposed to explain the behaviour of glass-forming materials. Nevertheless, we should remember that a strong correlation exists between volume and entropy, and that entropy-based theories of liquid dynamics<sup>9</sup> have been very successful<sup>10</sup>. Energy cannot be stored elastically under pressure without loss of entropy, so it is reasonable to expect a unified theory for fragile behaviour in soft matter to emerge in due course, thus bringing an explanation of the differing fragilities of liquids a step closer. ■

C. Austen Angell and Kazuhide Ueno are in the Department of Chemistry, Arizona State University, Tempe, Arizona 85287-1604, USA. e-mails: caa@asu.edu; kzueno@ynu.ac.jp

1. Mattsson, J. *et al.* *Nature* **462**, 83–86 (2009).
2. Bernal, J. D. *Nature* **183**, 141–147 (1959).
3. Woodcock, L. V. & Angell, C. A. *Phys. Rev. Lett.* **47**, 1129–1132 (1981).
4. Speedy, R. J. *J. Phys. Condens. Matter* **10**, 4185–4194 (1998).
5. Brito, C. & Wyart, M. *Europhys. Lett.* **76**, 149–155 (2006).
6. Cook, R. L., Herbst, C. A. & King, H. E. Jr *J. Phys. Chem.* **97**, 2355–2361 (1993).
7. Herbst, C. A., Cook, R. L. & King, H. E. Jr *J. Non-Cryst. Solids* **172–174**, 265–271 (1994).
8. Dyre, J. C. *Rev. Mod. Phys.* **78**, 953–972 (2006).
9. Adam, G. & Gibbs, J. H. *J. Chem. Phys.* **43**, 139–146 (1965).
10. Sastry, S. *Nature* **409**, 164–167 (2001).

## GENETICS

# Crossover control in two steps

Yonatan B. Tzur and Monica P. Colaiácovo

**During meiotic cell division, chromosome pairs exchange genetic material in a tightly controlled crossover process. Higher-order chromosome structure may regulate this genetic reshuffling at two distinct stages of meiosis.**

Meiosis is a cell-division program that reduces the chromosome complement by half, generating haploid gametes (eggs and sperm) from diploid germ cells. Meiosis also has another important function. It drives genetic diversity in a population by allowing genetic exchanges between chromosomes, in a sense 'reshuffling the chromosomal deck'. This genetic exchange takes place through the formation of double-strand breaks (DSBs) in the DNA which, when mended, can result in a crossover (CO) of genetic material between chromosome pairs (bivalents). Such COs are essential for generating physical attachment between chromosome pairs, which ensures that the sperm and eggs get the correct complement of chromosomes when the germ cell divides. CO formation is highly regulated, but how this process is controlled at a genome-wide or a region-specific level is not fully understood. A recent study published in *Cell* by Mets and Meyer<sup>1</sup> implicates higher-order chromosome structure in the regulation of the frequency and distribution of DSBs, and therefore of COs. The authors' analysis in the roundworm *Caenorhabditis elegans* also reveals that CO control operates in at least two temporally distinct stages of the meiotic recombination pathway.

After DSB formation in meiotic chromosomes, the broken ends of DNA are resected, creating 3' single-stranded DNA overhangs that are coated by the Rad51 recombination protein. This nucleoprotein filament invades an intact homologous sequence on the partner chromosome, creating a D-loop structure (Fig. 1). Capture of the second free DNA

end, DNA synthesis and ligation result in the formation of a recombination intermediate known as a double Holliday junction. This, when resolved, results in either the reciprocal exchange of genetic information between chromosomes (CO) or transfer of information to the repaired chromosome only (non-crossover, NCO)<sup>2</sup> (Fig. 1). Most NCOs, however, probably stem from an earlier repair mechanism termed synthesis-dependent strand annealing<sup>3</sup>.

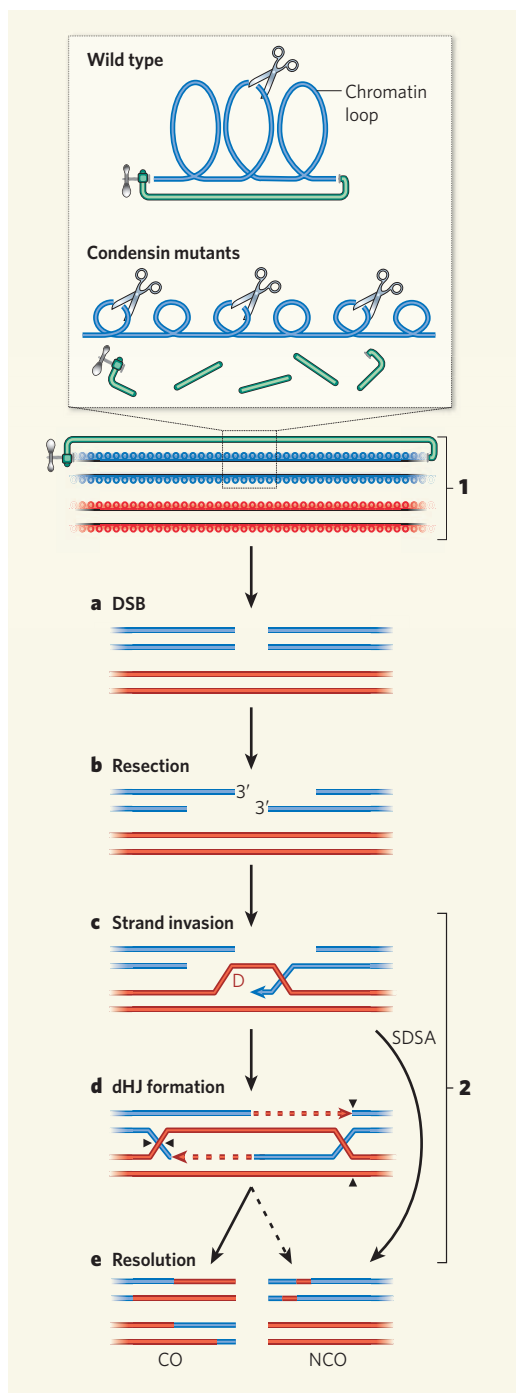
Studies in yeast, plants and other multicellular organisms<sup>3,4</sup> suggest that CO formation is controlled at several levels. First, COs are not evenly distributed along chromosomes: some genomic regions (recombination hotspots) are more prone to undergoing CO than others<sup>4,5</sup>. Second, each bivalent undergoes at least one CO (obligate CO)<sup>6</sup>. Third, a CO in one region of the chromosome decreases the likelihood that additional COs will occur in that vicinity (CO interference)<sup>7</sup>. Last, when DSB formation is reduced, the number of NCO events are sacrificed in favour of maintaining CO levels (CO homeostasis)<sup>8</sup>.

Mets and Meyer<sup>1</sup> take advantage of the extreme case of CO interference observed in *C. elegans* to investigate the mechanisms of CO regulation. Usually in this worm, only one CO occurs in each bivalent<sup>9</sup> and so a loss of CO interference is easy to detect. Previous work from the same group<sup>10</sup> had shown that the *C. elegans* protein DPY-28, a component of the dosage-compensation protein complex involved in the regulation of X-chromosome gene expression, also controls CO formation during meiosis. Mutations in the *dpy-28* gene increase the number of DSBs and the

frequency of COs, and reduce CO interference, so that more than one CO can take place per bivalent. Mets and Meyer's work<sup>1</sup>, as well as a study by Csankovszki and colleagues<sup>11</sup>, reveal three biochemically distinct condensin complexes in *C. elegans*: condensin I<sup>DC</sup>, involved in dosage compensation; condensin II, involved in mitotic and meiotic chromosome compaction; and a newly described condensin I complex, which includes DPY-28 and comprises five proteins in total, all of which are also found in either condensin I<sup>DC</sup> or condensin II (refs 12, 13). Similarly to mutations in *dpy-28*, mutations in the genes encoding any of the other four components of condensin I change the distribution of COs, and increase the number of double and triple COs in the *C. elegans* X chromosome. Moreover, this is not observed in a mutant that affects only the condensin I<sup>DC</sup> complex, and so reveals a specific function for the novel condensin I complex in CO regulation.

Mets and Meyer report that, in condensin-I-mutant worms, chromosome regions that have more COs have more DSBs (they measured these breaks by detecting the number of RAD-51 foci). Changes in the position of COs also correlate with changes in the sites of DSB formation. These results imply that condensin I might regulate CO formation on a chromosome-wide basis by affecting DSB production.

How might the condensin I complex affect DSB and CO formation? Chromosomes become condensed during meiosis, adopting an organization in which chromatin loops extend from the chromosome's axis or structural core<sup>14</sup>. Condensin protein complexes are a central contributor to this compacted organization<sup>15</sup>. In line with this, Mets and Meyer<sup>1</sup> observe that mutations in condensin I subunits increase the axis lengths of meiotic chromosomes (importantly, induction of DSBs with  $\gamma$ -irradiation does not). Mutations in subunits of the condensin II complex also increase the length of the chromosome axes, the number of DSBs and the amount of double COs. However,



**Figure 1 | Crossover formation and control.** **a**, During meiosis, chromosomes undergo double-strand break (DSB) formation. **b**, The broken ends of DNA are resected, producing 3' single-stranded DNA tails with which the Rad51 protein (not shown) associates. **c**, The single-stranded DNA invades a homologous DNA sequence on the chromosome pair, creating a D-loop structure. **d**, Capture of the second free DNA end, DNA synthesis and ligation result in a double Holliday junction (dHJ) intermediate structure. **e**, dHJs can be resolved by cleavage either as a crossover (CO) (if cleaved at the arrowheads shown in **d**) with reciprocal exchange of genetic information between chromosomes, or as a non-crossover (NCO) with transfer of genetic information to the repaired chromosome only. However, NCOs probably occur mostly through synthesis-dependent strand annealing (SDSA). Mets and Meyer<sup>1</sup> propose that, in the roundworm *Caenorhabditis elegans*, CO control can operate at two steps. The first step (1) is controlled by condensins I and II, protein complexes that compact chromosome structure during meiosis (green clamp). The inset shows how condensin protein complexes seem to regulate the number and distribution of DSBs (scissors), and hence of COs. The increase in chromosome axis length in condensin mutants (broken clamp) might reflect changes in the number and size of chromatin loops. This may, in turn, lead to changes in the frequency and distribution of DSBs (which are perhaps more common at loops), and consequently in changes in the number and distribution of COs. The second CO control step (2) may follow DSB formation, when the CO-versus-NCO decision is made.

hotspots, perhaps by exposing different loci to the DSB machinery. Given that DSBs have been proposed to occur preferentially at chromatin loops<sup>16</sup>, altering the density and position of loops could change the frequency and distribution of DSBs, and consequently COs (Fig. 1). This suggests that CO control — for example, ensuring the formation of the obligate CO — is exerted early in meiosis, either prior to or at DSB formation, and implicates condensins I and II in mediating this control on a chromosome-wide basis. Changes in chromosome structure might also create and relieve tension along the chromosome, which could be translated into the CO-versus-NCO decision, accounting, at least in part, for CO interference, as has been suggested previously<sup>17</sup>. However, the precise molecular basis for the execution of this decision remains to be defined. The effect of condensins and other elements that control chromosome structure may differ among individuals, perhaps because of varying levels of expression of these elements, and interplay with different cofactors. This could explain both the variance and dynamics of CO control.

But there are still outstanding questions. Are condensin complexes involved in CO control in organisms such as yeast and mammals? In these organisms, central features of meiosis differ from those of *C. elegans* — for example, more DSBs occur per chromosome, and CO interference is weaker. Are condensin complexes also involved in the control of CO homeostasis? How do the two distinct condensin complexes coordinate the CO control they exert? Continued studies of CO control in various organisms will be crucial to shedding more light on these questions.

Yonatan B. Tzur and Monica P. Colaiácovo are in the Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.  
e-mail: mcolaiacovo@genetics.med.harvard.edu

1. Mets, D. G. & Meyer, B. J. *Cell* **139**, 73–86 (2009).
2. Pâques, F. & Haber, J. E. *Microbiol. Mol. Biol. Rev.* **63**, 349–404 (1999).
3. Martinez-Perez, E. & Colaiácovo, M. P. *Curr. Opin. Genet. Dev.* **19**, 105–112 (2009).
4. Kauppi, L., Jeffreys, A. J. & Keeney, S. *Nature Rev. Genet.* **5**, 413–424 (2004).
5. Petes, T. D. *Nature Rev. Genet.* **2**, 360–369 (2001).
6. Jones, G. H. *Symp. Soc. Exp. Biol.* **38**, 293–320 (1984).
7. Muller, H. J. *Am. Nat.* **50**, 193–221 (1916).
8. Martini, E., Diaz, R. L., Hunter, N. & Keeney, S. *Cell* **126**, 285–295 (2006).
9. Meneely, P. M., Farago, A. F. & Kauffman, T. M. *Genetics* **162**, 1169–1177 (2002).
10. Tsai, C. J. *et al. Genes Dev.* **22**, 194–211 (2008).
11. Csanokovszki, G. *et al. Curr. Biol.* **19**, 9–19 (2009).
12. Chan, R. C., Severson, A. F. & Meyer, B. J. *J. Cell Biol.* **167**, 613–625 (2004).
13. Hagstrom, K. A., Holmes, V. F., Cozzarelli, N. R. & Meyer, B. J. *Genes Dev.* **16**, 729–742 (2002).
14. Zickler, D. & Kleckner, N. *Annu. Rev. Genet.* **33**, 603–754 (1999).
15. Hagstrom, K. A. & Meyer, B. J. *Nature Rev. Genet.* **4**, 520–534 (2003).
16. Blat, Y., Protacio, R. U., Hunter, N. & Kleckner, N. *Cell* **111**, 791–802 (2002).
17. Kleckner, N. *et al. Proc. Natl Acad. Sci. USA* **101**, 12592–12597 (2004).

the two condensin complexes probably work independently to control COs by controlling chromosome structure and hence the position and frequency of DSB formation.

In *C. elegans*, the authors found an average of 2.1 DSBs per meiotic chromosome pair, with a range of between 0 and 9. But the number of bivalents with no DSBs was markedly lower than expected if the distribution of breaks occurs at random (as predicted by a Poisson distribution). This implies that an active mechanism ensures that there is at least one DSB per bivalent — a prerequisite to achieve an obligate CO. In addition, most bivalents have about 2–6 DSBs, but only one CO, which suggests that CO control must also be enforced at a step after

DSB formation, during the CO-versus-NCO decision.

Many questions surround the process of meiotic CO control. Why and how do CO hotspots appear and disappear during population evolution? What is the molecular basis for the process controlling obligate CO and CO interference? Are these processes connected? If so, how is CO control executed at both the whole chromosome level and locally within short genomic intervals? Mets and Meyer's work<sup>1</sup> supports a model that starts to address some of these questions.

Changes in chromosome structure imposed by condensins I and II may alter both the frequency and the distribution of recombination





# 140<sup>TH</sup> BIRTHDAY MISCELLANY

*With this issue, it is 140 years since Nature first appeared on 4 November 1869. To mark the anniversary, these two pages offer a miscellany from that issue and from 1889, 1909, 1929, 1949, 1969 and 1989.*

POPPER/GETTY IMAGES



**4 November 1869.** Contributors to the first issue concern themselves with such topics as the fertilization of winter-flowering plants, meetings of "German naturalists and physicians at Innsbruck" and the American Philosophical Society, and with papers in *Zeitschrift für Chemie* and *Giornale di Scienze di Palermo*. Other news of the day features the opening of the Suez Canal, worry about the welfare of the explorer David Livingstone, and the following:

The "Female Physicians" question, thanks to Professor Masson, has made a great stride during the past week. Ladies are to be admitted to study Medicine at Edinburgh University. Imagine the feelings of the non-contents when Professor Masson, in a final outburst, described their argumentation as "rampageous mysticism, dashed with drivel from Anacreon!"

**7 November 1889.** Four pages in this issue are devoted to a dense but sweeping survey of the science of the preceding 20 years. In chemistry, the happy conclusion is that "the periodic law of Mendeleeff has become established as a generalization of the first importance". In the physical sciences, "The pregnant suggestion of Maxwell that light is an electro-magnetic phenomenon has borne good fruit". A comment on biology has a cryptic sting in the tail:

Apart from Darwinism, the most remarkable development of biological studies during these "twice ten tedious years" is undoubtedly the sudden rise and gigantic progress of our knowledge of the Bacteria ... We now know, through the labours of Toussaint, Chauveau, Pasteur, and Koch, of a number of diseases

which are definitely caused by Bacteria. We have also learnt from Pasteur how to control the attack of some of these dangerous parasites. Within these twenty years the antiseptic surgery founded by Sir Joseph Lister has received its full measure of trial and confirmation, whilst his opportunities and those of his fellow-countrymen for making further discoveries of a like kind have been ignorantly destroyed by an Act of Parliament.

**7 November 1889.** The twentieth anniversary issue could also find space for the following anecdote. The scene could have been drawn from a novel by Thomas Hardy (then in full creative flow; *Tess of the d'Urbervilles* was published in 1891).

During the terrific storm of the 12<sup>th</sup> of July last, a labourer's cottage was struck by lightning at Leagrave, near here. The lightning descended, according to an eye-witness's report, like a "spout of fire," and struck and descended the chimney, which it destroyed. In the room below there was an old shepherd, an invalid woman, a child and a shepherd's dog. The shepherd was sitting in a chair leaning on a stick, a kettle was boiling on the fire, and the door was open. The lightning entered the room simultaneously by the chimney and an adjoining window. The window was utterly destroyed, and the kettle was thrown from the fire across the room, the stick on which the shepherd was leaning was torn from his hand and also thrown across the room, the lightning entered a cupboard containing glass and crockery and destroyed every article, and plaster was torn from the walls. The man and woman remained unhurt, but the child was thrown down and its knees stiffened. The dog was struck perfectly stiff, "like a log of wood," and was considered dead. The room seemed full of fire, water, and sulphur, and the occupants said the smell of sulphur was so strong that they would certainly have been suffocated had it not been for the open door. After the storm had abated, the dog, with all its limbs stiff, was laid in a barn, where it very slowly and partially recovered. It long remained both deaf and blind, and was entirely dependent upon smell for its recognition of persons and things.



BETTMANN/CORBIS

**4 November 1909.** A 20-year-old Chandrasekhara Venkata Raman writes from India about vibrations in acoustics, one of his preoccupations. His greater fame rests on his subsequent discovery of a "new kind of radiation or light emission from atoms and molecules" — the Raman effect — for which he won a Nobel prize in 1930.

It is generally recognised that the nodes of a string which is maintained permanently in oscillation in two or more loops cannot be points of absolute rest, as the energy requisite for the maintenance of the vibrations is transmitted through these points. I have not, however, seen anywhere a discussion or experimental demonstration of some peculiar properties of this small motion. A brief note may therefore be of interest.

In the first place, the small motion at the node is in a phase which is different from that of the rest of the string. The exact difference of phase is shown by a dynamical investigation to be a quarter of an oscillation. The motion is of very small amplitude, and it

#### Image details

The Suez Canal, which opened 140 years ago; C. V. Raman, photographed in 1930; passengers board a Comet airliner in 1952; fission yeast.

#### History online

Full text of the first issue of *Nature* is available at [www.nature.com/nature/about/first](http://www.nature.com/nature/about/first)  
A brief illustrated history of the journal is at [www.nature.com/nature/history/index.html](http://www.nature.com/nature/history/index.html)

might therefore be thought a difficult matter to verify this experimentally. I have, however, devised some convenient arrangements with which this can be effected. I shall here mention only one method: this was to compound the oscillation at every point on the string with another perpendicular to it of half the frequency, and to observe the compound oscillation at the nodes and elsewhere.

Such a compound oscillation can easily be maintained permanently by having the string attached to the prong of an electrically maintained tuning-fork, so that it lies in a plane perpendicular to the prongs, but in a direction inclined to the line of their vibration. When the load on the string is slightly greater than that necessary for the most vigorous maintenance, points on the string describe parabolic arcs with concavities in opposite directions in alternate loops, the whole forming a beautiful and interesting type of stationary vibration. This is not, however, the stage convenient for observing the small motion at the nodes. When the tension of the string is relaxed, so as to make its vibration stronger, points on the string, i.e. except the node, describe 8 curves. The curve described by the node is neither a straight line nor an 8 curve, but is a very flat parabola. From this, the phase-relation between the small motion at the nodes and the large motion elsewhere is obvious ...

C. V. Raman

**2 November 1929.** "Our Astronomical Column" was an enduring feature, with reports such as this:

*La Science Moderne* for September contains an article on Mars by A. Nodon, president of the Astronomical Society of Bordeaux. He reproduces several drawings made by E. M. Antoniadi at Meudon and J. Comas Sola at Barcelona. A description of the temperature measures of Menzel, Coblentz, and Lampland leads to the conclusion that the conditions are likely to be suitable for the presence of vegetation, and that the various tints observed in the darker regions of the disc are compatible with this view of their nature. Some of the drawings show numbers of small round dark markings, the positions of which appear to change between one opposition and the next. It is noted that snow or frost on the ground can be distinguished from cloud or mist in the air by the greater permanence in the position of the former. There are some speculations about the future of the earth; it is suggested that in the distant future the ocean may have largely disappeared, and the height of the mountains have been greatly diminished by denudation, which would bring about a resemblance to the present condition of Mars.



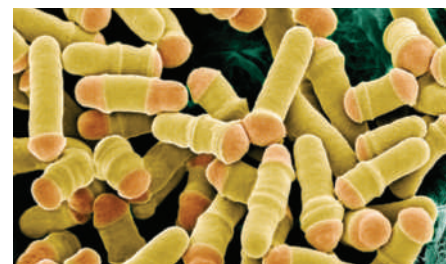
**5 November 1949.** The era of mass air-transport dawns with the Comet airliner. After early success, a problem with metal fatigue caused the plane to be grounded and undergo redesign before re-entering service. In updated form, a later incarnation of the aircraft is the Nimrod.

A de Havilland 'Comet' jet-propelled air liner recently carried out a long-distance flight, as part of its trials, from London to Castel Benito, Tripoli, a distance of 2,980 miles in a total flying time of six hours thirty-eight minutes — an average speed of just over 450 miles an hour. This is more than 100 miles an hour faster than the authenticated speed of any established air liner at present in use. It left London at 6:30 a.m. and, after a two-hour period in Tripoli for refuelling, it arrived back in London at 4 p.m. A performance such as this suggests that this type of machine should be taken up immediately by air transport concerns everywhere. Quantity production is already in hand by Messrs. de Havilland; but it is not expected that they will be operating on the trunk services of the British Overseas Airways Corporation until 1952–53.

**1 November 1969.** An editorial spares no sarcasm in commenting on the unseemly scramble to ban cyclamates in Britain.

The cyclamate bandwagon has made farcical progress in the past week, though it is hard to say whether scientific advisers or the politicians who manipulated them look the more ridiculous. Within a week of the United States ban on the artificial sweetener, at least five other countries had followed suit, in several instances before their advisers had seen the new evidence. From the precipitate unction with which politicians in Sweden, Finland, Japan, Canada and the United Kingdom proclaimed their renunciation of the artificial sweetener, it might have been supposed that cyclamates presented some kind of hazard to health. If the new evidence had shown that they transmuted into arsenic at the drop of a hat, the substances could not have been banned more quickly ... The postscript to this seven day wonder has been afforded by a spokesman at the Ministry of Agriculture. Asked why such terrible speed

was necessary, he admitted that "Public opinion wanted an answer straight away. We'd have been caught very much with our head in the sand if we didn't do anything about the cyclamates. We had either to ban them or prove them harmless". Ostrich-like or not, even the fleet footed Ministry of Agriculture could not prove cyclamates harmless at seven days' notice and, discretion being the better part of valour, it only remained for housewives to be soothed with advice as to what foods and beverages were likely to contain the deadly substance before a grateful public could sink back, probably to soothe itself with a well earned cigarette, confident in the knowledge that its political and scientific guardians would leave no stone unturned to protect the nation's health, save any such as might be inconveniently heavy to uproot.



**2 November 1989.** Research on the cell cycle is in highly fruitful mode. A paper by Kathleen L. Gould and Paul Nurse, of which the following is the title, was part of the stream of discovery that culminated in a Nobel prize shared by Leland H. Hartwell, Tim Hunt and Nurse in 2001.

#### Tyrosine phosphorylation of the fission yeast *cdc2<sup>+</sup>* protein kinase regulates entry into mitosis

K. L. Gould & P. Nurse, ICRF Cell Cycle Group, Microbiology Unit, Department of Biochemistry, University of Oxford

**2 November 1989.** Virginia Trimble brings data to bear on an enduring fact of academic life.

For unto him that hath ... The concept of P-selection, by which graduates of the more prestigious (P) universities tend to publish longer than those from other (NP) universities, is herewith introduced into the literature. If you really want a career in scientific research, then get your PhD from the most prestigious university that will take you ... To be precise, astronomers who received their doctoral degrees from a university that ranks repeatedly among the 'top three' American astronomical PhD programmes continue to publish in the field about twice as long as those with degrees from a university that ranks in the 'second ten'.



# Cooperation between non-kin in animal societies

Tim Clutton-Brock<sup>1</sup>

**Explanations of cooperation between non-kin in animal societies often suggest that individuals exchange resources or services and that cooperation is maintained by reciprocity. But do cooperative interactions between unrelated individuals in non-human animals really resemble exchanges or are they a consequence of simpler mechanisms? Firm evidence of reciprocity in animal societies is rare and many examples of cooperation between non-kin probably represent cases of intra-specific mutualism or manipulation.**

As Darwin appreciated, cooperative behaviour—actions adapted to assist others that involve costs to the fitness of participants—poses a fundamental problem to the traditional theory of natural selection, which rests on the assumption that individuals compete to survive and breed<sup>1</sup>. Early explanations of cooperative behaviour in animal societies (see Figs 1–4) often argued that it was maintained by benefits to groups or populations<sup>2,3</sup>, but most recent theoretical treatments are rooted in Hamilton's models of the evolution of cooperation based on the concept of inclusive fitness<sup>4</sup> and it is now widely accepted that selection operating through benefits to non-descendant kin is commonly involved in maintaining cooperative behaviour (see Box 1).

## Alternative paths to cooperation between non-kin

Although kin selection theory provides a satisfactory explanation of cooperation between kin, cooperation between unrelated individuals remains a problem and the evolutionary mechanisms that maintain it are still debated. The most frequent suggestion is that individuals exchange resources or services, suffering temporary net costs as a result of providing assistance, which are exceeded by subsequent benefits when they subsequently receive assistance from individuals that they have previously helped<sup>5,6</sup>. Explanations of this kind are derived from Trivers' concept of reciprocal altruism<sup>7</sup>, although they are now usually referred to as examples of 'direct' or 'cost-counting' reciprocity to avoid the implication that they involve altruism.

In his original description of reciprocal altruism, Trivers<sup>7</sup> interpreted cooperative interactions between members of different species (as well as between members of the same species) as reciprocal exchanges of assistance, and pointed out their similarity to the Prisoner's Dilemma games of economists<sup>8</sup>. In these games, two individuals that are constrained to interacting with each other initially agree to cooperate and gain higher pay-offs if they do so than if they refuse, although the highest pay-off is gained by individuals that defect when their partners cooperate. In iterated versions, cooperation can persist if individuals assist each other in turn and avoid partners that are unlikely to reciprocate<sup>5,6</sup>. The most successful strategies are slightly 'generous' ones where individuals copy the previous behaviour of their partners, cooperating when they do and responding to defection by ceasing to provide assistance, but forgiving occasional lapses<sup>6</sup>. An additional tactic that can reduce the chance that individuals will assist non-cooperators is to monitor interactions between third parties, and models of 'indirect reciprocity' have explored the effects of mechanisms of this

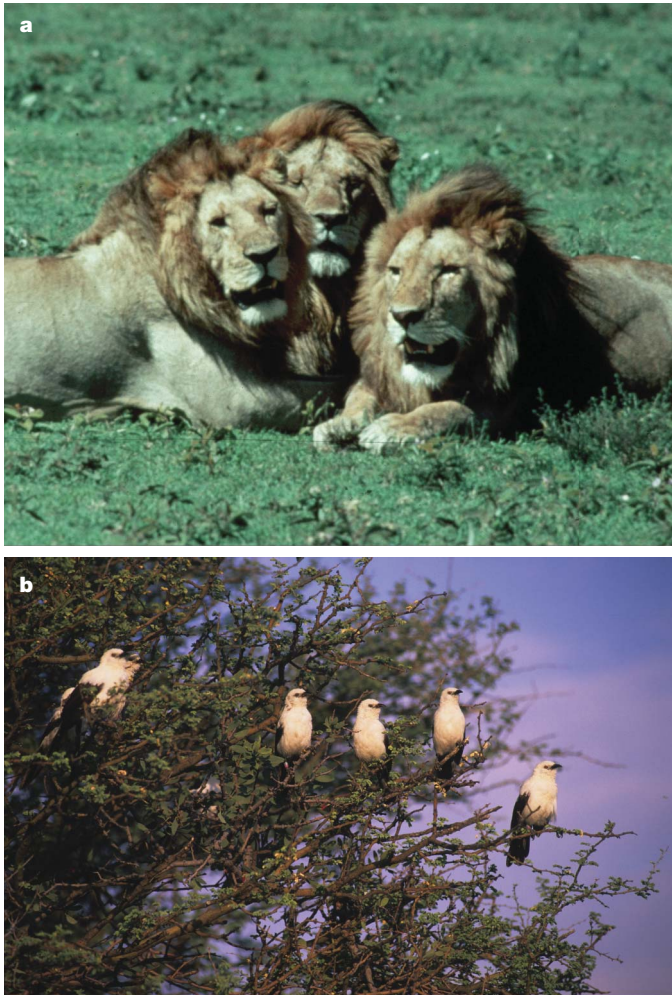
kind<sup>9–11</sup>. In addition, recent models of 'generalized reciprocity' show that, if cooperative individuals tend to associate with each other, cooperation can be stable if individuals respond cooperatively after they have been assisted by another group member<sup>12,13</sup>, irrespective of their identity.

Although reciprocity is often suggested as an explanation of cooperation between non-kin, several other evolutionary mechanisms are also capable of maintaining cooperation between unrelated individuals. In some cases, cooperation generates immediate synergistic benefits shared by cooperators that exceed the costs of providing assistance<sup>14,15</sup> as in the 'public goods' games of economists<sup>16</sup>. Mechanisms of this kind probably maintain many mutualistic interactions between species, like the hunting associations found between raptors and carnivores<sup>17</sup> or between different predatory fish<sup>18</sup>, as well as many examples of cooperative foraging between unrelated conspecifics, like the cooperative manoeuvres of foraging pelicans<sup>19</sup>. I refer to them as cases of mutualism, though some prefer to use the term only to refer to cooperation between members of different species<sup>20</sup>.

Alternatively, cooperative interactions may be a consequence of manipulative strategies—cases in which the behaviour of one (or both) partners is adapted solely to maximising its own immediate fitness. For example, dominant individuals commonly use coercive tactics (including harassment and punishment) to force others to provide assistance at some cost to their fitness<sup>21,22</sup>. Manipulation can also involve actions that increase the fitness of their partners. In some cases, manipulators may adjust their behaviour to increase the probability that their partners perform purely selfish actions which increase their own fitness but have coincidental, 'by-product' benefits to the manipulator—a form of manipulation known as 'pseudo-reciprocity' that is probably common in animal strategies<sup>23,24</sup>. In others, they may use inducements to exploit the fixed tendencies of others, such as their tendency to habituate to regular neighbours (see later). Finally, it has been suggested that cooperative behaviour may represent a costly display that signals an individual's prowess as a rival or its potential quality as a mate<sup>25</sup>, though as yet there is little evidence that this is the case<sup>26</sup>.

There is a contrast in the assumptions of models of reciprocity and models of mutualism and manipulation. Whereas models of reciprocity assume that providing assistance has net costs at the time that it is provided, which are offset by subsequent benefits, models of mutualism and manipulation usually assume that the benefits of assistance exceed the costs involved at the time that it is provided. A consequence

<sup>1</sup>Department of Zoology, University of Cambridge, Downing Street, Cambridge CB3 9EJ, UK.



**Figure 1 | Territorial choruses in birds and mammals.** **a**, Territorial choruses by male lions discourage intruders and are likely to benefit all contributors. (The photograph was taken by C. R. Packer.) **b**, The territorial choruses of pied babblers and many other group-living birds also serve to demarcate territories and to deter intruders. (The photograph was taken by T.H.C.-B.)

of this difference is that, in reciprocity models, cheating strategies that exploit the time-lag between cost and benefit are often favoured, whereas the assumptions of explanations based on mutualism and manipulation usually exclude this possibility. Which of the two approaches to accounting the fitness benefits of providing assistance is appropriate depends less on the actual timing of effects (given that most actions only affect breeding success or survival hours, days or months after they are performed) than on the inevitability of net benefits to cooperators. Where the benefits of providing assistance are not inevitable, so that selection can favour cheating strategies that



**Figure 2 | Cooperative mobbing of potential predators by meerkats drives away potential predators.** (The photograph was taken by T.H.C.-B.)

exploit the behaviour of cooperators, it is realistic to treat fitness benefits as deferred. However, where providing assistance generates unavoidable benefits that cannot be exploited by cheating strategies, the method of accounting used in models of mutualism and manipulation may be more appropriate. Box 2 provides examples of both scenarios. In the first case, the benefits of cooperation are guaranteed by the synchrony of costs and benefits but, in other cases, they may be maintained by asymmetries in cost or benefit between partners or by contrasts in the constraints operating on different individuals (see below).

Over the last thirty years, theoretical studies of cooperation have focused principally on developing models of reciprocity between non-kin<sup>6</sup> while empirical studies have explored the role of kinship and, to a lesser extent, of mutualistic interactions and manipulative strategies<sup>27–29</sup>. The need for an integrated explanation of cooperative behaviour has recently led to systematic attempts to define key concepts and to classify models and mechanisms capable of maintaining cooperative behaviour between non-kin<sup>20,30–32</sup>. For example, recent reviews<sup>31</sup> suggest successive distinctions (1) between cases where actions benefit others but involve no costs to initiators and actions involving costs to initiators, (2) between cases where benefits are extracted or voluntary, and (3) between cases where the responses of recipients are costly or cost-free. However, cooperative interactions often involve more than one evolutionary mechanism so that examples of discrete categories of cooperation are rare, and there is the danger that detailed classifications provoke semantic arguments and focus disproportionate attention on unusual mechanisms<sup>33</sup>.

The most immediate need is now to assess the relative importance of broader categories of explanations for the maintenance of cooperation in animal societies. How common *is* cooperation between non-kin? Does it commonly involve temporary fitness costs? Do cooperative interactions between non-kin really approximate to exchanges of resources or services as models of the Prisoner's Dilemma suggest, or



**Figure 3 | Cooperative hunting in African wild dogs.** The hunting and breeding success of pack members increases with group size. (The photograph was taken by R. Woodroffe.)





**Figure 4 | Alternating bouts of allo-grooming in social primates are one of the best documented examples of reciprocity.** (The photograph was taken by T.H.C.-B.)

is cooperation more commonly maintained by immediate shared benefits? Do unrelated animals often assist each other as a consequence of manipulative tactics involving coercion or inducement? Answers to these three questions are still unresolved and are of central importance to the direction of future research as well as to the interpretation of results of existing studies. The next three sections evaluate

### Box 1 | Cooperating with kin

Where animals form stable breeding groups the members of which assist each other, individuals of one (or, in some cases, both) sexes are usually related to each other. To explain the evolution of cooperation between relatives, Hamilton<sup>4</sup> introduced the concept of 'inclusive fitness', consisting of the 'direct' fitness individuals derive from producing descendants and the 'indirect' fitness that they derive from helping non-descendant relatives minus any benefits received from them<sup>27,71,87,88</sup>. Most recent theoretical treatments are rooted in Hamilton's models of the evolution of cooperation based on the concept of inclusive fitness<sup>4</sup>: this can be reduced to the proposition that cooperation should evolve wherever  $Br - C > 0$  where  $r$  is the coefficient of relatedness between two individuals,  $B$  is the benefit of cooperation to the recipient and  $C$  is the cost to the co-operator (Hamilton's Rule).

Over the last forty years, empirical studies of cooperation have provided extensive evidence of the importance of indirect or 'kin-selected' benefits in maintaining cooperative behaviour in animals<sup>68,89–92</sup>. Among both invertebrates and vertebrates, cooperative behaviour between unrelated individuals that are not potential breeding partners is seldom highly developed even if they live in stable groups or colonies. Most of the more costly forms of cooperative behaviour, such as the provisioning of young born to other individuals, are restricted to species living in groups consisting primarily of relatives<sup>59,68</sup>. For example, cooperative or eusocial breeding appears to be restricted to lineages with monandrous mating systems, which generate relatively high levels of relatedness<sup>93,94</sup>. Where the sexes differ in their tendency to associate with relatives, cooperation is typically more frequent and more highly developed in whichever sex associates with relatives<sup>59</sup>. In species where individuals have to choose between assisting close relatives, distant relatives or unrelated individuals, they typically show a strong preference for joining and assisting close relatives<sup>71,95,96</sup>, unless some form of coercion is involved. And where group members compete, competition between non-kin is usually more intense than between kin<sup>97</sup>.

### Box 2 | The Soldier's Dilemma

The contrast between explanations of cooperation based on immediate shared benefits and those based on reciprocity can be illustrated by analogy with the pay-offs of different strategies to ambushed soldiers. Suppose that a four-man patrol of soldiers is ambushed by an unseen number of opponents, that soldiers who fire back attract incoming fire and increase the absolute chance that they will be killed from zero to 10%, and that the probability that the patrol will be overrun (and all members killed) is 100% if no individuals fire back but declines by 25% for each individual that does fire back. In this case, even if all three other members of a patrol fire back, the fourth member can still increase his chances of survival by 15% by doing so too. Under these conditions there is no dilemma: the best strategy for surviving a tour of duty is to fire back when attacked. Cooperative interactions of this kind are common in animal societies and often involve non-kin as well as kin<sup>59</sup>. For example, the synchronized roaring of male lions faced by intruders<sup>98,99</sup> (Fig. 1a) may resemble the first situation: roaring is likely to be relatively cheap while the potential costs of a successful and determined intrusion by bachelor males to all group members may be very large. Territorial choruses in birds (Fig. 1b) and synchronized mobbing of predators (Fig. 2) may also be maintained by similar net benefits.

Alternatively, suppose that only two members of the ambushed patrol closest to the enemy are in imminent danger and that their chances of surviving depend on the two furthest from the ambush providing covering fire. Assuming that the same individuals support each other in successive patrols, cooperation (firing back) can be maintained through some form of reciprocity<sup>5,6</sup>. However, in contrast to the first scenario, the temporal separation of the costs of providing assistance and the benefits when the favour is eventually returned provides an opportunity for cheats to exploit the delay, unless indirect reciprocity or the threat of punishment causes the costs of cheating to exceed the costs of cooperation. Cooperative interactions in which individuals do not gain automatic benefits from providing assistance do occur in animal societies but they more usually involve kin or potential breeding partners. For example, social mongooses will attempt to rescue group members caught by dangerous predators in coordinated attacks on the predator<sup>100</sup>. Similar interactions involving unrelated individuals are common in human societies in which cheats are usually punished, censured or excluded<sup>83,101</sup>.

the importance of reciprocity, mutualism and manipulation in maintaining cooperative behaviour between non-kin in animal societies.

### Reciprocity

Following the publication of Trivers' paper in 1971 (ref. 7), empirical studies provided a wide range of potential examples of direct reciprocity, including reciprocal assistance in mating competition in olive baboons *Papio anubis*<sup>34</sup>, the exchange of blood meals in vampire bats *Desmodus rotundus*<sup>35</sup> and sequential predator inspection by fish<sup>36</sup>. Their results stimulated thinking about how variation in the costs and benefits of cooperation might affect the behaviour of individuals, and led to the development of 'market models' predicting how variation in the supply of services or in the benefits they confer may affect how much individuals are ready to 'pay' for them<sup>37</sup>. The language of human exchanges—negotiate, trade, barter, swap, payment, debt, commodity, currency, settlement, market—came to be widely used to describe cooperative interactions between animals<sup>38</sup>. In addition, studies suggested that animals often 'interchange' one type of resource or service for another<sup>39</sup> and pay for access to resources or services. For example, a recent study of macaques which showed that males were more likely to mate with females they have previously groomed was immediately interpreted as evidence of prostitution in animals<sup>40,41</sup>. More recently, dissatisfaction with models of the Prisoner's Dilemma as a conceptual framework for interpreting cooperative behaviour in animal societies has increased and several reviews have pointed out contrasts between the assumptions of these models and the conditions under which animals interact<sup>20,42</sup>. Asymmetries in power and status between cooperating

individuals are often pronounced, generating large differences in the value of assistance which affect the form of interactions and relationships<sup>37</sup>. Partners often communicate their intentions and modify their behaviour in the course of repeated interactions<sup>43,44</sup>. Unlike prisoners, social animals are seldom constrained to cooperate with particular partners and can develop profitable relationships and terminate unproductive ones<sup>45</sup>. Partners that are short-changed may punish defectors or join rival coalitions, so that cheating may have substantial costs<sup>21,22</sup>. While there are certainly analogies between cooperative behaviour in animals and the interactions of partners in Prisoner's Dilemma games or human exchanges, these are not as close as is frequently implied.

Dissatisfaction with empirical evidence of direct reciprocity, especially in natural populations, has also grown<sup>28,31,32,42</sup>. To provide convincing evidence that cooperative interactions are maintained by direct reciprocity, empirical studies need to demonstrate that the same individuals assist each other repeatedly; that the frequency with which they give assistance is contingent on the frequency with which they receive it; that cooperative behaviour has temporary net costs to the cooperator's fitness and is adapted to provide benefits to their partners; and that partners are not close relatives or prospective mates. Although experimented studies with captive birds and mammals have shown that the probability that individuals will assist each other can be affected by the previous behaviour of partners or other group members<sup>46–48</sup>, very few attempts have been made to measure the net fitness benefits of cooperation between non-kin in natural populations and it is seldom clear whether cooperative actions entail temporary net costs or immediate net benefits. As a result, few of the studies commonly cited as examples of direct reciprocity in natural populations provide evidence that can definitely exclude the possibility that cooperative behaviour is maintained by immediate shared benefits, manipulative tactics or kin selection (Table 1). For example, although the same male baboons cooperate with each other to steal mating partners from rival males<sup>34</sup>, there is no evidence that each male benefits in turn, and a simpler alternative is that males gain immediate benefits by synchronising their attacks on rivals<sup>49</sup>. Similarly, while blood sharing in vampire bats is widely cited as an example of reciprocity<sup>50</sup>, field studies do not provide unequivocal evidence that individuals are more likely to give blood to bats that previously have provided them with a meal<sup>51</sup>. Several other explanations are feasible: begging may prevent well-fed bats from resting, so that it may pay individuals to donate some of their resources to begging neighbours (see ref. 52) or, alternatively, a proportion of group members are relatives<sup>51</sup> and blood sharing may normally involve kin.

The cooperative interactions between non-kin that most closely resemble the exchanges envisaged by models of direct reciprocity are

those in which two or more individuals repeatedly assist each other within narrow time frames. For example, in many social mammals, allo-grooming sessions in which partners groom each other in turn often involve unrelated individuals as well as relatives<sup>53</sup>. Across grooming sessions, the duration of bouts given by each partner is usually similar and grooming is usually interpreted as an example of reciprocity<sup>54,55</sup>. However, in many interactions of this kind, selection may favour individuals that prolong sessions until they have received a fixed amount of assistance (in this case, grooming) to avoid the costs or risks associated with attempting to find an alternative partner, and both partners may parcel their assistance in bouts to minimize the risk of defection. So, although these interactions share some of the properties of the Prisoner's Dilemma models, opportunities for repeated cheating are limited and an alternative interpretation is that they represent examples of pseudo-reciprocity<sup>56</sup>.

Mutualism

Long before the publication of Hamilton's papers<sup>4</sup>, philosophers and scientists (including Darwin) stressed the potential importance of shared, mutualistic benefits in maintaining cooperative behaviour in animal societies<sup>1,2</sup>. More recent studies provide convincing empirical evidence of effects of this kind and show that cooperation can generate immediate shared benefits. For example, cooperation between hunting partners can increase their per capita success in catching or defending food<sup>57</sup> (Fig. 3). Similarly, where groups defend communal territories and inter-group competition is frequent, larger groups commonly displace smaller ones, gradually eroding the size of their territories and depriving their members of access to resources<sup>28,58</sup>. And, in species where group members rear young communally, the (per-capita) costs of raising young to breeders and helpers often decline with rising group size and their success increases<sup>28,59</sup>. Mutualistic benefits may also help to maintain the long-term affiliative relationships between individuals that are found in many social animals<sup>60,61</sup>. In some social primates, individuals form long-lasting relationships with non-kin involving association and mutual grooming<sup>62,63</sup> and are more likely to support other animals in competitive interactions if they have recently been groomed by them<sup>64,65</sup>. Analysis of long-term data show that differences in the extent and quality of affiliative relationships can affect breeding success<sup>66</sup>. Although both partners must gain from these relationships for it to be worth maintaining them, these relationships may be more realistically interpreted as mutualistic coalitions or alliances than as examples of direct reciprocity.

A common objection to explanations of cooperative behaviour based on shared benefits is that situations of this kind generate opportunities for selection to favour individuals that free-ride on

Table 1 | A personal evaluation of examples of direct reciprocity

Proposed reciprocity	Presence of convincing empirical evidence from natural populations that:				
	The same individuals repeatedly assist each other	Cooperative behaviour is adapted to provide benefits to their partners	The frequency of contributions by individuals is adjusted to the frequency of their partners' contributions	Recipients of assistance do not include relatives or potential mates	Assistance has net fitness costs at the time it is provided in natural populations
Interspecific cleaning in fish	Yes	No	Yes	Yes	No
Egg trading in hermaphroditic fish	Yes	Yes	Yes	No	No
Predator inspection in fish	Yes	No	Yes	Yes	No
Tolerance of floaters in birds	Yes	No	No	Yes	No
Food calling in birds	No	Yes	Yes	Yes	No
Alarm calls in birds	No	Yes	No	No	No
Grooming in ungulates	Yes	Yes	Yes	Yes	No
Alarm signals in deer	Yes	Yes	No	No	No
Allo-parental care in carnivores	No	Yes	No	No	No
Midwifery in bats	No	Yes	No	No	No
Blood sharing in vampire bats	Yes	Yes	No	No	No
Allo-grooming in primates	Yes	Yes	Yes	Yes	No
Grooming interchanges in primates	Yes	Yes	Yes	Yes	No
Supportive coalitions in male baboons	Yes	No	No	Yes	No
Mutual support between female primates	Yes	Yes	No	Yes	No
Food-sharing in chimpanzees	No	No	No	Yes	No

Potential examples of direct reciprocity are taken from ref. 68.



the activities of cooperators, preventing the maintenance of cooperation<sup>67,68</sup>. This need not necessarily be the case if group size is relatively small and cooperation generates inevitable benefits to the cooperator's breeding success or survival, so that cheats cannot exploit the delay between giving and receiving assistance and there is little opportunity for free-riding strategies to evolve<sup>14</sup> (see Box 2). Many examples of cooperation between non-kin in animal societies, including cooperative foraging or hunting, cooperative defence of territories and mates and cooperative construction of nests and burrows, are likely to generate immediate net benefits which cannot easily be exploited by cheats. Moreover, although shared benefits are capable of maintaining cooperation between non-relatives, in practice, interactions commonly occur in stable, kin-based groups where the indirect benefits of cooperation are likely to reduce the relative benefits of cheating strategies<sup>14</sup>.

Where cooperative behaviour is maintained by shared direct benefits, variation in fitness can, in theory, be partitioned into variation between groups and variation between group members<sup>69</sup>. For example, in the first situation described in Box 2, it would be possible to partition fitness into within-group effects generated by the costs of retaliation to individuals and between-group effects generated by differences in the proportion of individuals that retaliate. On these grounds, some theorists prefer to regard cooperative behaviour as a product of group selection<sup>69</sup> but it is often more practical to compare the relative pay-offs of different strategies to individuals in the population as a whole, treating cooperation maintained by shared benefits as a product of selection operating through variation in the inclusive fitness of individuals<sup>32,70,71</sup>.

## Manipulation

Manipulative interactions differ from the exchanges of resources or services envisaged by models of direct reciprocity in several ways: the behaviour of both individuals need not be adapted to providing assistance to their partner; the costs and benefits of interactions often differ substantially between partners so that there is no alternation of assistance given and received; and cooperative actions or inducements are likely to generate immediate net benefits rather than net costs. In many animal societies, individuals use coercive tactics to manipulate others to perform actions likely to increase their fitness or to avoid activities that jeopardise their interests<sup>21,22</sup>. In some cases, individuals harass other group members until they assist them: for example, chimpanzees will harass individuals that have killed monkeys until they are given a share of the kill<sup>52</sup>. Alternatively, dominant individuals may punish subordinates that fail to assist them: for example, dominant rhesus monkeys punish subordinates that fail to give contact calls when they discover good feeding sites<sup>72</sup>. In both cases, coercion is likely to force individuals to adopt strategies that would not be optimal if their behaviour was unconstrained.

Manipulation need not involve coercion and can have benefits to other group members as well as costs. For example, in many social animals, the purely selfish activities of individuals can generate coincidental, unselected benefits to other group members known as 'by-product mutualisms'<sup>27,68</sup>. Where these have substantial effects on fitness, selection may favour individuals that adjust their behaviour to maximise the by-product benefits that they receive, generating examples of 'pseudo-reciprocity'<sup>24,73</sup>. For example, some eusocial insects protect or provision symbionts whose numbers contribute to their breeding success or survival<sup>74</sup>. In other cases, manipulators modify their behaviour to take advantage of the fixed responses of conspecifics: for example, by regularly associating with dominant individuals, and grooming them repeatedly, subordinates may habituate them to their presence, gaining shelter from competition and interference by rank neighbours<sup>39,75,76</sup> (Fig. 4). Similarly, in many primate societies, non-breeding females are strongly attracted by infants and may gain benefits (ranging from experience of parenting to the facilitation of future alliances) by interacting with them<sup>77</sup>. Mothers are initially protective of their infants, but non-breeding females modify their behaviour to induce mothers to allow them

access to their infant, sitting close to them, signalling their friendly intentions and grooming them, so that mothers gradually habituate to their presence and become less restrictive, and persistent 'suitors' are able to hold or play with their infants<sup>77</sup>.

Manipulative tactics also have a role in the establishment and maintenance of long-term mutualistic relationships. In many social mammals that live in stable groups, individuals compete to establish relationships with potential protectors, allies or mates, using a wide range of different forms of affiliative behaviour, including close association, grooming, support in competitive interactions, reassurance and consolation<sup>60,61,78</sup>. Where alliances are disrupted by conflicts, they may attempt to minimize the consequences of conflicts by reconciling with competitors, supporting winners or consoling losers<sup>79–81</sup>, and where their status is threatened by rival alliances, they may attempt to disrupt relationships between their rivals by strategic cooperation with their allies<sup>60</sup>.

In many social animals, attempts to establish and maintain useful relationships with other individuals represent one component of a more general strategy of modifying their social environment so as to maximise their fitness. In some species, this involves the use of coercive or cooperative tactics to facilitate the development of long-lasting affiliative relationships with dominant animals, potential allies or prospective mates<sup>60–62</sup>. In others, similar behaviour is used to adjust the size, age structure or kin composition of the group they live in<sup>59</sup>. In such cases, attempts to assess the costs and benefits of cooperative actions within narrow time frames or to treat interactions as isolated games are likely to be misleading.

## Discussion

Research over the last thirty years shows that cooperation in animal societies most frequently involves kin and is seldom highly developed in groups consisting of unrelated individuals (see Box 1). In many cases where non-kin do cooperate with each other in natural populations, one or both partners seem likely to gain immediate benefits from their behaviour and evidence that cooperation has temporary fitness costs is rare. As a result future studies need to be cautious in interpreting cooperative interactions between non-kin as examples of reciprocity.

Theoretical research now needs to explore the evolution of mutualistic interactions and manipulative tactics with as much energy as it has previously devoted to investigating the role of reciprocity. Many interactions between non-kin are likely to be maintained by a combination of mutualism, coercion and inducement and we need to know more about how the different processes that can maintain cooperation interact with each other. In addition, while it is convenient to consider the evolution of mutualistic and manipulative behaviour in non-kin, these strategies are most highly developed in stable groups which typically consist predominantly of relatives, so that both direct and indirect fitness benefits are likely to be involved. Kinship might be expected to affect the readiness of individuals to share benefits and to manipulate each other's activities but, as yet, we know little about these interactions.

The results of research on cooperation in animal societies emphasize both the contrasts and the similarities with human behaviour. Like other primates, humans often assist relatives, share benefits and use coercion and inducement to manipulate each other. As in other social animals, the size of groups may often have an important influence on the fitness of individuals<sup>82,83</sup>. However, in contrast to other species, reciprocal exchanges of costly services or valuable resources between non-kin are widespread and often involve considerable time delays between assistance given and received and extensive opportunities for cheating<sup>16,84</sup>. A likely reason for this contrast between humans and other primates is that reliable exchanges of valuable resources require some form of language (and associated psychological capacities) to establish the intentions and expectations of both parties regarding the nature and timing of exchanges<sup>84,85</sup> as well as the social norms that discourage cheating<sup>83,86</sup>. Lacking the ability to make specific agreements about

future events, other animals may commonly be restricted to cooperative strategies that generate immediate benefits to their inclusive fitness.

1. Darwin, C. *The Origin of Species* Ch. VIII (The Modern Library, 1859/1958). Although Darwin did not confront the problem of altruistic behaviour directly, he was aware of the challenge posed by the development of sterile castes in some social insects and, in Chapter VIII of *The Origin of Species*, he describes how he thought, at first, that this was fatal to the whole theory of natural selection.
2. Kropotkin, P. *Mutual Aid* (William Heineman, 1908).
3. Wynne-Edwards, V. C. *Animal Dispersion in Relation to Social Behaviour* (Oliver and Boyd, 1962).
4. Hamilton, W. D. The genetical evolution of social behaviour. I. *J. Theor. Biol.* **7**, 1–52 (1964).  
Although they are usually cited as an explanation of the role of kin selection in maintaining cooperative behaviour, Hamilton's papers and Hamilton's Rule (see Box 1) also provide the basis for explaining cooperation between non-kin.
5. Axelrod, R. *The Evolution of Cooperation* (Basic Books, 1984).
6. Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1565 (2006).
7. Trivers, R. L. The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57 (1971). The theoretical interest generated by Trivers' theory caused explanations of cooperation between non-kin based on reciprocity to eclipse all other explanations for more than thirty years.
8. von Neumann, J. & Morgenstern, O. *Theory of Games and Economic Behaviour* (Princeton Univ. Press, 1953).
9. Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity. *Nature* **437**, 1295–1297 (2005).
10. Suzuki, S. & Akiyama, E. Evolution of indirect reciprocity in groups of various sizes and comparison with direct reciprocity. *J. Theor. Biol.* **245**, 539–552 (2007).
11. Roberts, G. Evolution of direct and indirect reciprocity. *Proc. R. Soc. Lond. B* **275**, 173–179 (2008).
12. Pfeiffer, T., Rutte, C., Killingback, T., Taborsky, M. & Benhoeffer, S. Evolution of cooperation by generalized reciprocity. *Proc. R. Soc. B* **272**, 1115–1120 (2005).
13. Rankin, D. J. & Taborsky, M. Assortment and the evolution of generalized reciprocity. *Evolution* **63**, 1913–1922 (2009).
14. Kokko, H., Johnstone, R. A. & Clutton-Brock, T. H. The evolution of cooperative breeding through group augmentation. *Proc. R. Soc. Lond. B* **268**, 187–196 (2001). Kokko's model shows that shared fitness benefits can maintain costly cooperation between non-kin in small groups—but that their capacity to do so is greatly enhanced if  $r$  exceeds 0.
15. Brown, J. S. & Vincent, T. L. Evolution of cooperation with shared costs and benefits. *Proc. R. Soc. Lond. B* **275**, 1985–1994 (2008).
16. Bowles, S. & Gintis, H. in *Genetic and Cultural Evolution of Cooperation* (ed. Hammerstein, P.) 429–444 (Dahlem Univ. Press, 2003).
17. Paxton, M. Foraging associations between pale chanting goshawks, honey badgers and slender mongooses. *Gabar* **3**, 82–93 (1988).  
As Trivers pointed out, cooperative interactions between members of different species cannot be attributed to kin selection—but this does not necessarily suggest that they represent examples of reciprocity.
18. Bshary, R., Hohner, A., Ait-el-Djoudi, K. & Fricke, H. Interspecific communicative and coordinated hunting between groupers and giant Moray eels in the Red Sea. *PLoS Biol.* **4**, 2393–2398 (2006).
19. McMahon, B. F. & Evans, R. M. Foraging strategies of American white pelicans. *Behaviour* **120**, 69–89 (1992).
20. West, S. A., Griffin, A. S. & Gardiner, A. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.* **20**, 415–432 (2007).  
West *et al.* prefer to use mutualism to refer only to interactions between species, whereas I use it here to refer to all  $+/+$  interactions, where the behaviour of individuals is adapted to generating benefits to their partners, whether they involve heterospecifics or conspecifics.
21. Clutton-Brock, T. H. & Parker, G. A. Punishment in animal societies. *Nature* **373**, 209–216 (1995).
22. Cant, M. A. & Johnstone, R. A. Self-serving punishment and the evolution of cooperation. *Evol. Biol.* **19**, 1383–1385 (2006).
23. Connor, R. C. Pseudoreciprocity: investing in mutualism. *Anim. Behav.* **34**, 1652–1654 (1986).
24. Leimar, O. & Connor, R. C. in *Genetic and Cultural Evolution of Cooperation* (ed. Hammerstein, P.) 202–222 (MIT Press, 2003).  
The models of pseudo-reciprocity produced by Connor and Leimar have attracted less attention than models of reciprocity but the processes they capture may help to explain many examples of cooperation between non-kin in animals.
25. Zahavi, A. & Zahavi, A. *The Handicap Principle: a Missing Piece of Darwin's Puzzle* (Oxford Univ. Press, 1997).
26. Wright, J., Berg, E., de Kort, S. R., Khazin, V. & Maklakov, A. A. Cooperative sentinel behaviour in the Arabian babbler. *Animal Behav.* **61**, doi:10.1006/anbe.2001.1838 (2001).
27. Brown, J. L. *Helping and Communal Breeding in Birds* (Princeton Univ. Press, 1987).
28. Clutton-Brock, T. H. Breeding together: kin selection and mutualism in cooperative vertebrates. *Science* **296**, 69–72 (2002).
29. Russell, A. F. in *Ecology and Evolution of Cooperative Breeding in Birds* (eds Koenig, W. & Dickinson, J.) 210–227 (Cambridge Univ. Press, 2004).
30. Lehmann, L. & Keller, L. The evolution of cooperation and altruism: a general framework and classification of models. *J. Evol. Biol.* **19**, 1365–1378 (2006).
31. Bergmüller, R., Russell, A. F., Johnstone, R. A. & Bshary, R. Integrating cooperative breeding into theoretical concepts of cooperation. *Behav. Process.* **76**, 61–72 (2007).  
Bergmüller's review provides a comprehensive classification of types of reciprocity.
32. West, S. A., Griffin, A. S. & Gardner, A. Evolutionary explanations for cooperation. *Curr. Biol.* **17**, R661–R672 (2007).
33. Taborsky, M. Cooperation built the Tower of Babel. *Behav. Process.* **76**, 95–99 (2007).
34. Packer, C. Reciprocal altruism in olive baboons (*Papio anubis*). *Nature* **265**, 441–443 (1977).
35. Wilkinson, G. S. Reciprocal food sharing in the vampire bats. *Nature* **308**, 181–184 (1984).
36. Milinski, M. No alternative to Tit for Tat in sticklebacks. *Anim. Behav.* **39**, 989–991 (1990).
37. Noë, R. & Hammerstein, P. Biological markets: supply and demand determine the effect of partner choice in cooperation, mutualism and mating. *Behav. Ecol. Sociobiol.* **35**, 1–11 (1994).  
Market theory provides important predictions concerning the magnitude of costs and benefits that apply as much to explanations of cooperation based on mutualistic interactions or manipulative tactics as to those based on reciprocity, so evidence that its predictions are confirmed does not necessarily support the interpretation of cooperative interactions as exchanges.
38. de Waal, F. B. M. & Berger, M. L. Payment for labour in monkeys. *Nature* **404**, 563 (2000).
39. Watts, D. P. Reciprocity and interchange in the social relationships of wild male chimpanzees. *Behaviour* **139**, 343–370 (2002).
40. Gummert, M. D. Payment for sex in a macaque mating market. *Anim. Behav.* **74**, 1655–1667 (2007).
41. Barras, C. A biological market for monkey prostitution. *New Sci.* **197**, 6 (2008).  
This article represents an extreme example of the tendency to emphasise similarities between cooperative interactions between non-kin in animals and exchanges in human societies—but there are many others.
42. Hammerstein, P. in *Genetic and Cultural Evolution of Cooperation* (ed. Hammerstein, P.) 83–93 (Berlin Univ. Press, 2003).
43. Smuts, B. B. & Watanabe, J. M. Social relationships and ritualized greetings in adult male baboons *Papio cynocephalus*. *Int. J. Primatol.* **11**, 147–172 (1990).
44. McNamara, J. M., Gasson, I. E. & Houston, A. I. Incorporating rules for responding to evolutionary games. *Nature* **401**, 368–371 (1999).
45. Enquist, M. & Leimar, O. Evolution of cooperation in mobile organisms. *Anim. Behav.* **45**, 747–757 (1993).
46. Clements, K. C. & Stephens, D. W. Testing models of non-kin cooperation: mutualism and the prisoner's dilemma. *Anim. Behav.* **50**, 527–535 (1995).
47. Hauser, M. D., Chen, M. K. & Chen, F. Give unto others: genetically unrelated cotton-top tamarin monkeys preferentially give food to those who altruistically give food back. *Proc. R. Soc. Lond. B* **270**, 2363–2370 (2003).
48. Rutte, C. & Taborsky, M. The influence of social experience on cooperative behaviour of rats (*Rattus norvegicus*): direct versus generalised reciprocity. *Behav. Ecol. Sociobiol.* **62**, 499–505 (2008).  
Rutte and Taborsky provide convincing evidence that prior experience of cooperative interactions in rats affects their tendency to respond cooperatively to other individuals.
49. Bercovitch, F. B. in *Reproductive Ecology and Human Evolution* (ed. Ellis, P. T.) 369–396 (Aldine de Gruyter, 2001).
50. Dawkins, R. *The God Delusion* (Bantam, 2006).
51. Wilkinson, G. S. Reciprocal altruism in bats and other mammals. *Ethol. Sociobiol.* **9**, 85–100 (1988).  
Wilkinson's study of blood sharing in vampire bats is commonly cited as the classic example of reciprocity but his results do not exclude the possibility that cooperation is maintained by kin selection, mutualism or manipulation.
52. Gilby, I. C. Meat sharing among the Gombe chimpanzees: harassment and reciprocal exchange. *Anim. Behav.* **71**, 953–963 (2006).
53. Hart, B. L. & Hart, L. Reciprocal allogrooming in impala. *Anim. Behav.* **44**, 1073–1083 (1992).
54. Barrett, L., Henzi, S. P., Weingrill, T., Lycett, J. E. & Hill, R. A. Market forces predict grooming reciprocity in female baboons. *Proc. R. Soc. Lond. B* **266**, 665–670 (1999).  
Barrett and Henzi's studies of grooming in wild baboons provide some of the most sophisticated evidence that grooming affects the extent of tolerance and the frequency of aggression in natural populations, but can be interpreted as examples of pseudo-reciprocity (or manipulation) as well as of evidence of reciprocity.
55. Barrett, L. & Henzi, S. P. in *Cooperation in Primates and Humans* (eds Kappeler, P. M. & Schaik, C. P. v.) 209–232 (Springer, 2006).
56. Connor, R. C. Impala allogrooming and the parceling model of reciprocity. *Anim. Behav.* **49**, 528–530 (1995).
57. Creel, S. & Creel, N. M. *The African Wild Dog: behavior, ecology and conservation* (Princeton Univ. Press, 2001).



58. Wrangham, R. W. An ecological model of female-bonded primate groups. *Behaviour* **75**, 262–300 (1980).
59. Clutton-Brock, T. H. in *Cooperation in Primates and Humans* (eds Kappeler, P. M. & van Schaik, C. P.) 173–190 (Springer, 2006).
60. de Waal, F. B. M. *Chimpanzee Politics* (Allen and Unwin, 1982).
61. Smuts, B. B. *Sex and Friendship in Baboons* (Aldine, 1985).
62. Mitani, J. C. Male chimpanzees form enduring and equitable social bonds. *Anim. Behav.* **77**, 633–640 (2009).
63. Gomes, C. M., Mundry, R. & Boesch, C. Long-term reciprocation of grooming in wild West African chimpanzees. *Proc. R. Soc. Lond. B* **276**, 699–706 (2009).
64. Seyfarth, R. M. & Cheney, D. L. Grooming alliances and reciprocal altruism in vervet monkeys. *Nature* **308**, 541–543 (1984).
65. Hemelrijk, C. K. Support for being groomed in long-tailed macaques *Macaca fascicularis*. *Anim. Behav.* **48**, 479–481 (1994).
66. Silk, J. B., Alberts, S. C. & Altmann, J. Social bonds of female baboons enhance infant survival. *Science* **302**, 1331–1334 (2003).
67. Boyd, R., & Richerson, P. J. The evolution of reciprocity in sizeable groups. *J. Theor. Biol.* **132**, 337–356 (1988).
68. Dugatkin, L. A. *Cooperation Among Animals: An Evolutionary Perspective* (Oxford Univ. Press, 1997).
- Dugatkin's comprehensive review of cooperative behaviour in animals provides an important synthesis of theory and evidence.**
69. Wilson, D. S. & Wilson, E. O. Rethinking the theoretical foundation of sociobiology. *Q. Rev. Biol.* **82**, 327–348 (2007).
70. West, S. A., Griffin, A. S. & Gardiner, A. Social semantics: how useful has group selection been? *J. Evol. Biol.* **21**, 374–385 (2007b).
71. Gardiner, A. & Grafen, A. Capturing the superorganism: a formal theory of group adaptation. *J. Evol. Biol.* **22**, 659–671 (2009).
72. Hauser, M. D. Costs of deception: cheaters are punished in rhesus monkeys *Macaque mulatta*. *Proc. Natl Acad. Sci. USA* **89**, 12137–12139 (1992).
73. Connor, R. C. Altruism among non-relatives: alternatives to the 'prisoner's dilemma'. *Trends Ecol. Evol.* **10**, 84–86 (1995).
74. Bourke, A. F. G. in *Behavioural Ecology: an Evolutionary Approach* (eds Krebs, J. R. & Davies, N. B.) 203–227 (Blackwell, 1997).
75. Barrett, L., Gaynor, D. & Henzi, S. P. A dynamic interaction between aggression and grooming reciprocity among female chacma baboons. *Anim. Behav.* **63**, 1047–1053 (2002).
76. Cheney, D. L. & Seyfarth, R. M. *Baboon Metaphysics: the Evolution of a Social Mind* (Univ. Chicago Press, 2007).
77. Henzi, S. P. & Barrett, L. Infants as a commodity in a baboon market. *Anim. Behav.* **63**, 915–921 (2002).
78. de Waal, F. B. M. The integration of dominance and social bonding in primates. *Q. Rev. Biol.* **61**, 459–479 (1986).
79. Aureli, F., van Schaik, C. P. & van Hoof, J. Functional aspects of reconciliation among captive long-tailed macaques. *Am. J. Primatol.* **19**, 39–51 (1989).
80. Aureli, F. Post-conflict behaviour among wild long-tailed macaques *Macaca fascicularis*. *Behav. Ecol. Sociobiol.* **31**, 329–337 (1992).
81. Seed, A. M., Clayton, N. S. & Emery, N. J. Postconflict third-party affiliation in rooks, *Corvus frugilegus*. *Curr. Biol.* **17**, 1–7 (2007).
82. Turchin, P. *War and Peace and War: the Rise and Fall of Empires* (Pi Press, 2005).
83. Richerson, P. J. & Boyd, R. *Not By Genes Alone: How Culture Transformed Human Evolution* (Univ. Chicago Press, 2005).
84. Smith, E. A. in *Genetic and Cultural Evolution of Cooperation* (ed. Hammerstein, P.) 401–428 (Dahlem Univ. Press, 2003).
85. Stevens, J. R. & Hauser, M. D. Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn. Sci.* **8**, 60–65 (2004).
86. Boyd, R., Gintis, H., Bowles, S. & Richerson, P. J. The evolution of altruistic punishment. *Proc. Natl Acad. Sci. USA* **100**, 3531–3535 (2003).
87. Grafen, A. How not to measure inclusive fitness. *Nature* **298**, 425–426 (1982).
88. Grafen, A. in *Behavioural Ecology: an Evolutionary Approach* (eds Krebs, J. R. & Davies, N. B.) 62–84 (Blackwell, 1984).
89. Trivers, R. L. & Hare, H. Haplo-diploidy and the evolution of the social insects. *Science* **191**, 249–263 (1976).
90. Bourke, A. F. G. & Franks, N. R. *Social Evolution in Ants* (Princeton Univ. Press, 1995).
91. Solomon, N. G. & French, J. A. *Cooperative Breeding in Mammals* (Cambridge Univ. Press, 1997).
92. Koenig, W. & Dickinson, J. (eds) *Ecology and Evolution of Cooperative Breeding in Birds* (Cambridge Univ. Press, 2004).
93. Boomsma, J. J. Kin selection versus sexual selection: why the ends do not meet. *Curr. Biol.* **17**, R673–R683 (2007).
94. Hughes, W. O. H., Oldroyd, B. P., Beekman, M. & Ratnieks, F. L. W. Ancestral monogamy shows kin selection is key to the evolution of eusociality. *Science* **320**, 1213–1216 (2008).
95. Emlen, S. T. & Wrege, P. H. The role of kinship in helping decisions among white-fronted bee-eaters. *Behav. Ecol. Sociobiol.* **23**, 305–315 (1988).
96. Hatchwell, B. J., Ross, D. J., Fowlie, M. K. & McGowan, A. Kin discrimination in cooperatively breeding long-tailed tits. *Proc. R. Soc. Lond. B* **268**, 885–890 (2001b).
97. Boncoraglio, G. & Saino, N. Barn swallow chicks beg more loudly when brood mates are unrelated. *J. Evol. Biol.* **21**, 256–262 (2007).
98. Packer, C. *et al.* in *Reproductive Success* (ed. Clutton-Brock, T. H.) 363–383 (Univ. Chicago Press, 1988).
99. Heinsohn, R. G. & Packer, C. Complex cooperative strategies in group-territorial African lions. *Science* **269**, 1260–1262 (1995).
100. Rood, J. P. Banded mongoose rescues back member from eagle. *Anim. Behav.* **31**, 1261–1262 (1983).
101. Young, P. in *Genetic and Cultural Evolution of Cooperation* (ed. Hammerstein, P.) (Dahlem Univ. Press, 2003).

**Acknowledgements** I thank M. Bell, S. Bowles, R. Boyd, M. Cant, A. Cockburn, N. Davies, A. Gardner, S. Hodge, R. Johnstone, A. Russell, S. Sharp, J. Silk, J. Stevens, M. Taborsky and S. West for discussions or for comments on this manuscript. I also thank C. R. Packer and R. Woodroffe for the use of their photographs.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence should be addressed to the author (thcb@cam.ac.uk).

## ARTICLES

# An oestrogen-receptor- $\alpha$ -bound human chromatin interactome

Melissa J. Fullwood<sup>1</sup>, Mei Hui Liu<sup>1</sup>, You Fu Pan<sup>1</sup>, Jun Liu<sup>1</sup>, Han Xu<sup>1</sup>, Yusoff Bin Mohamed<sup>1</sup>, Yuriy L. Orlov<sup>1</sup>, Stoyan Velkov<sup>1</sup>, Andrea Ho<sup>1</sup>, Poh Huay Mei<sup>1</sup>, Elaine G. Y. Chew<sup>1</sup>, Phillips Yao Hui Huang<sup>1</sup>, Willem-Jan Welboren<sup>2</sup>, Yuyuan Han<sup>1</sup>, Hong Sain Ooi<sup>1</sup>, Pramila N. Ariyaratne<sup>1</sup>, Vinsensius B. Vega<sup>1</sup>, Yanquan Luo<sup>1</sup>, Peck Yean Tan<sup>1</sup>, Pei Ye Choy<sup>1</sup>, K. D. Senali Abayratna Wansa<sup>1</sup>, Bing Zhao<sup>1</sup>, Kar Sian Lim<sup>1</sup>, Shi Chi Leow<sup>1</sup>, Jit Sin Yow<sup>1</sup>, Roy Joseph<sup>1</sup>, Haixia Li<sup>1</sup>, Kartiki V. Desai<sup>1</sup>, Jane S. Thomsen<sup>1</sup>, Yew Kok Lee<sup>1</sup>, R. Krishna Murthy Karuturi<sup>1</sup>, Thoreau Herve<sup>1</sup>, Guillaume Bourque<sup>1</sup>, Hendrik G. Stunnenberg<sup>2</sup>, Xiaolan Ruan<sup>1</sup>, Valere Cacheux-Rataboul<sup>1</sup>, Wing-Kin Sung<sup>1,3</sup>, Edison T. Liu<sup>1</sup>, Chia-Lin Wei<sup>1</sup>, Edwin Cheung<sup>1,4,5</sup> & Yijun Ruan<sup>1,4</sup>

Genomes are organized into high-level three-dimensional structures, and DNA elements separated by long genomic distances can in principle interact functionally. Many transcription factors bind to regulatory DNA elements distant from gene promoters. Although distal binding sites have been shown to regulate transcription by long-range chromatin interactions at a few loci, chromatin interactions and their impact on transcription regulation have not been investigated in a genome-wide manner. Here we describe the development of a new strategy, chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) for the *de novo* detection of global chromatin interactions, with which we have comprehensively mapped the chromatin interaction network bound by oestrogen receptor  $\alpha$  (ER- $\alpha$ ) in the human genome. We found that most high-confidence remote ER- $\alpha$ -binding sites are anchored at gene promoters through long-range chromatin interactions, suggesting that ER- $\alpha$  functions by extensive chromatin looping to bring genes together for coordinated transcriptional regulation. We propose that chromatin interactions constitute a primary mechanism for regulating transcription in mammalian genomes.

Although genomic information is usually presented as a linear series of bases, genomes are known to be organized into three-dimensional structures *in vivo* through interactions with protein factors for nuclear process such as transcription<sup>1</sup>. The precise and coordinated regulation of transcription requires the binding of transcription factors to specific regulatory DNA sequences in the genome. Chromatin immunoprecipitation (ChIP) microarray<sup>2</sup> (ChIP-Chip) and ChIP-sequencing<sup>3,4</sup> (ChIP-PET and ChIP-Seq) have identified global transcription-factor-binding sites (TFBSs) and revealed that many TFBSs are far from gene promoters<sup>5</sup>. For example, most TFBSs bound by ER- $\alpha$  in the human genome are distal to transcription start sites (TSSs) of target genes<sup>6–10</sup>. A major question arising from this observation is which distal TFBSs are non-functional fortuitous binding sites, and which are involved in transcriptional activity through a remote control mechanism. Long-range chromatin interactions between DNA elements engaged in transcriptional regulation<sup>11,12</sup> have been observed with the use of chromosome conformation capture (3C)<sup>13,14</sup> and variants including ChIP-3C<sup>15,16</sup>, 4C<sup>17–20</sup>, 5C<sup>21</sup> and 6C<sup>22</sup>, and also RNA tagging and recovery of associated proteins (RNA TRAP)<sup>23</sup> and fluorescence *in situ* hybridization (FISH)<sup>24</sup>. However, these methods are limited to one-point or partial genome-oriented detection and are incapable of the *de novo* detection of genome-wide chromatin interactions<sup>25</sup>.

To address whether and how DNA elements bound by protein factors interact through long-range chromatin looping in a genome-wide and unbiased manner, we conceived a new strategy, which we called ChIA-PET. We applied ChIA-PET to characterize ER- $\alpha$ -bound

chromatin interactions in oestrogen-treated human breast adenocarcinoma cells (MCF-7), and generated the first human chromatin interactome map. (A ChIA-PET visualization browser is provided at <http://cms1.gis.a-star.edu.sg> (username 'guest', password 'gisimsgtb') for viewing the ER- $\alpha$  ChIA-PET map.) Furthermore, using active promoter and transcriptional marks such as trimethylation of lysine 4 on histone H3 (H3K4me3) and RNA polymerase II (RNAPII) from ChIP sequencing as well as gene expression microarray data, we show that ER- $\alpha$ -bound chromatin interactions are functionally involved in regulating specific genes.

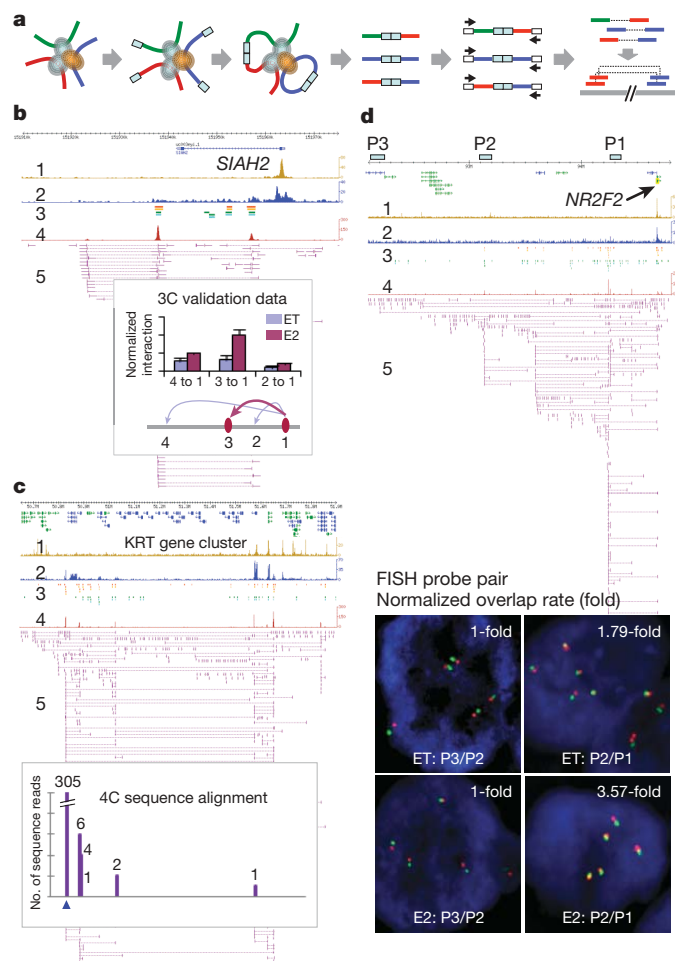
## The ChIA-PET method

In ChIA-PET, long-range chromatin interactions are captured by crosslinking with formaldehyde. Sonicated DNA–protein complexes are enriched by ChIP. Tethered DNA fragments in each of the chromatin complexes are connected with DNA linkers by proximity ligation, and paired-end tags (PETs) are extracted for sequencing. The resulting ChIA-PET sequences are mapped to reference genomes to reveal relationships between remote chromosomal regions brought together into close spatial proximity by protein factors (Fig. 1a and Supplementary Fig. 1).

ChIA-PET proximity ligation generates two types of ligation products: self-ligation of the same DNA fragments and inter-ligation between different DNA fragments. PET sequences derived from self-ligation products are mapped in the reference genome within a 3-kilobase (kb) span, demarcating ChIP DNA fragments, similar to

<sup>1</sup>Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore 138672. <sup>2</sup>Department of Molecular Biology, Nijmegen Centre for Molecular Life Sciences, Radboud University, 6500 HB Nijmegen, The Netherlands. <sup>3</sup>Department of Computer Science, School of Computing, National University of Singapore, Singapore 117543. <sup>4</sup>Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597. <sup>5</sup>School of Biological Sciences, Nanyang Technological University, Singapore 637551.





**Figure 1 | ChIA-PET method with validations.** **a**, ChIA-PET diagram. DNA fragments in sonicated, ChIP-enriched chromatin complexes were processed by linker ligation, proximity ligation, PET extraction, sequencing, and mapping to reveal interacting loci. **b**, ChIA-PET browser tracks: 1, H3K4me3 ChIP-Seq; 2, RNAPII ChIP-Seq; 3, ER- $\alpha$  (orange) and FoxA1 ChIP-chip (green)<sup>9</sup>; 4, ER- $\alpha$  ChIA-PET density; 5, inter-ligation PETs. Inset: 3C validation of interacting ER- $\alpha$ BSs (purple) and controls (blue) under ethanol control (ET) and induction with oestrogen (E2). Error bars show s.e.m. **c**, 4C validation, showing 4C bait region (blue) and interaction targets (purple bars) with counts of sequence hits. **d**, Validation by FISH, showing increased P2–P1 interactions under E2 induction with background normalization (P3/P2). FISH probe genomic locations (P1, P2 and P3) are indicated.

the standard ChIP-sequencing method<sup>3,8</sup>. Tethered DNA fragments in individual chromatin complexes can also ligate with each other, and the mapping results of such inter-ligation PET sequences would reveal if they are intrachromosomal (both tags of each PET are from the same chromosome) or interchromosomal (the tags are from different chromosomes). Singleton PETs are presumed experimental noise, and overlapping PET clusters are considered enriched putative binding sites or interaction events (Supplementary Fig. 2).

To test the ChIA-PET strategy, we constructed two ChIA-PET libraries from independent ER- $\alpha$  ChIP-enriched oestrogen-treated MCF-7 chromatin preparations, and generated two replicate pilot datasets (IHM001H and IHM001N) using Roche/454 pyrosequencing. Our analysis showed that both ChIA-PET libraries produced comparable putative binding sites and interactions. To assess levels of false positive chromatin interactions, we created a negative control ChIP-PET library (IHM043) from the same ChIP sample, wherein the DNA was reverse crosslinked before proximity ligation. We also analysed a previously reported cloning-based ChIP-PET library (SHC007)<sup>8</sup>. Both libraries generated abundant binding sites but no interactions. As an additional control, we used IgG, which binds to chromatin

nonspecifically, to perform a mock ChIA-PET analysis (IHM062), and only a few binding sites and interactions were identified (Table 1, Supplementary Figs 2 and 3 and Supplementary Note 1).

In proximity ligation-based analyses including 3C, the level of non-specific chimaeric DNA ligations between different chromatin complexes can be high and thus may confound data analysis. To address this, we designed linker nucleotide barcodes in the ChIA-PET method to specifically identify such chimaeric ligation PETs in another ER- $\alpha$  ChIA-PET replicate. Linker barcoding analysis suggests that chimaeric ligations are random and do not overlap with each other to form false positive interactions (Table 1, Supplementary Fig. 4 and Supplementary Note 2). A possible complication is that ChIP-enriched loci with more DNA fragments would result in proportionally higher chances of inter-ligations, leading to false-positive interactions comprising randomly overlapping inter-ligation PETs among highly-enriched ChIP DNA fragments. Hence, we devised a statistical scheme to calculate such probabilities and neutralize the potential ChIP-enrichment bias (Supplementary Methods; validations are given in Supplementary Fig. 5).

Together, these libraries indicate that the prevalent chromatin interactions (Supplementary Fig. 2d–g) identified by ER- $\alpha$  ChIA-PET data depend on proximity ligations of chromatin complexes and are neither technical artefacts of ligations between random DNA fragments nor mapping errors.

### ER- $\alpha$ -bound chromatin interactome map

Next we generated a large ER- $\alpha$  ChIA-PET dataset (IHM001F) with  $3.2 \times 10^7$  PET sequences by Illumina GAII paired-end sequencing (Table 1 and Supplementary Methods) for comprehensive analyses of ER- $\alpha$  binding and chromatin interactions in oestrogen-treated MCF-7 cells. Of  $4.6 \times 10^6$  uniquely mapped PET sequences,  $1.2 \times 10^6$  (27%) were self-ligation PETs. Among the self-ligation PETs, 16.7% clustered to form overlapping PET groups, representing 14,468 putative ER- $\alpha$ -binding sites (ER- $\alpha$ BSs) (false discovery rate (FDR) < 0.01, PET count per ER- $\alpha$ BS at least 5; Supplementary Table 1). Of the inter-ligation PETs,  $2.3 \times 10^5$  (5.1% of uniquely aligned PETs) were intrachromosomal and  $3.2 \times 10^6$  (68%) were interchromosomal (Table 1). After statistical analyses in which we discarded singleton inter-ligation PETs as either very weak interactions or background noise, defined clusters from overlapping inter-ligation PETs, corrected for ChIP enrichment biases and filtered out obviously false interactions due to structural variations in the MCF-7 genome (Supplementary Methods), we identified a large set of 1,451 intrachromosomal and a small set of 15 interchromosomal overlapping clusters consisting of three or more inter-ligation PETs per cluster (FDR < 0.05). These represent paired inter-ligating ChIP DNA fragments, which indicate potential distant chromatin interactions bound by ER- $\alpha$  (Supplementary Table 2).

Each chromatin interaction detected by an inter-ligation PET cluster features two anchor regions (interacting loci) and a loop (the intermediate genomic span between the two anchors), and is therefore called a ‘duplex interaction’ (Supplementary Table 2). Most anchors (1,893 out of 2,008 = 94%) involve self-ligation PET-defined ER- $\alpha$ BSs (FDR < 0.01). Many nearby duplex interactions are interconnected, linking three or more anchors into ‘daisy-chain’ aggregated complex interactions (Fig. 1b–d and Supplementary Fig. 6). For example, multiple duplex interactions with three ER- $\alpha$ BSs in the *SIAH2* region interconnect to form a complex interaction. Hence, we further assembled 1,036 duplex interactions into 274 complex interactions based on overlapping of interaction anchors (Supplementary Methods). The remaining interactions (415) were stand-alone duplex interactions. In all, we identified 689 ER- $\alpha$ -bound chromatin interaction regions (Supplementary Table 3).

To verify the ChIA-PET results, we validated several new ER- $\alpha$ BSs identified in this study by ChIP-quantitative polymerase chain reaction (qPCR) (Supplementary Fig. 7), as well as putative intrachromosomal interaction sites (20 genomic loci) by 3C, ChIP-3C, 4C and FISH experiments (three examples are shown in Fig. 1; others are

**Table 1 | Summary statistics of library PET sequences**

Library code	Library identity	Total PET	Unique PET	Self-ligation		Intrachromosome inter-ligation		Interchromosome inter-ligation	
				PET	PET clusters*	PET	PET clusters†	PET	PET clusters†
Small-scale testing of the ChIA-PET method									
IHM001N	ChIA-PET	715,369	271,648	78,706	2,701	16,677	176	176,265	0
IHM001H	ChIA-PET	764,899	293,754	103,740	3,405	17,718	215	172,296	0
IHM043	ChIP-PET	1,118,509	745,251	634,993	1,158	7,386	2‡	102,872	1
SHC007	ChIP-PET	361,241	214,668	192,511	489	2,196	0	19,961	0
IHM062	ChIA-PET (IgG)	436,248	217,708	40,847	0	11,254	0	165,607	0
Analysis of chimaeras									
IHH015M	ChIA-PET (AA + BB)	4,246,429	2,049,719	953,384	3,909	129,492	2,183	966,843	3
IHH015C	ChIA-PET (chimaeras)	5,904,476	1,790,714	15,490	35	98,805	0	1,676,419	0
Large-scale ChIA-PET analysis									
IHM001F	ChIA-PET	31,828,194	4,638,633	1,249,081	14,560	234,400	1,451	3,155,152	15
IHH015F	ChIA-PET	19,590,581	6,125,099	1,841,684	6,665	348,057	3,543	3,935,358	4

ChIA-PET data mapped at satellites and structural variation sites were removed.  
\* Self-ligation PET clusters for identifying binding sites (FDR < 0.01, PET count at least 5).  
† Inter-ligation PET clusters for identifying interactions include at least two (small-scale) or three (chimaeras and large-scale analysis) overlapping PETs (FDR < 0.05). Interchromosomal interactions were subjected to manual curation.  
‡ One interaction has a genomic span of less than 5 kb, suggesting that it results from extra-long self-ligation PETs, and the other has a genomic span of more than 10 Mb and PET counts of only 2, and so could be non-specific.

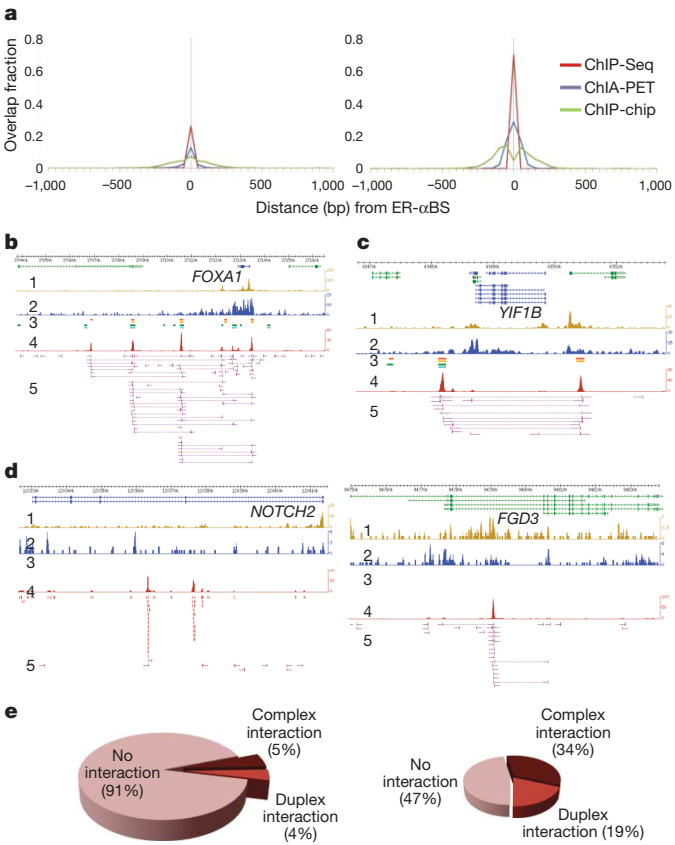
shown in Supplementary Figs 8–11 and Supplementary Tables 4 and 5). Moreover, the 3C and FISH experiments showed higher levels of chromatin interactions in oestrogen-treated conditions than in untreated conditions, indicating that the interactions are oestrogen-dependent. We also examined three putative interchromosomal interactions by FISH; however, none of them were positive (Supplementary Table 4 and Supplementary Note 3), suggesting that most ER- $\alpha$ -bound intra-chromosomal interactions were genuine, whereas the putative inter-chromosomal interactions were false positives or were too weak to be validated.

Taken together, the ER- $\alpha$ BS and chromatin interactions identified by ChIA-PET data constitute a whole-genome chromatin interaction map bound by ER- $\alpha$ . The genomic spans of most duplex interactions (86%) are less than 100 kb, about 13% are from 100 kb to 1 megabase (Mb), and less than 1% are more than 1 Mb. Complex interactions extend genomic span by connecting multiple duplex interactions. Many complex interactions (47%) have genomic spans in the range 100 kb to 1 Mb, with a few that are more than 1 Mb (Supplementary Fig. 12 and Supplementary Table 3).

To determine the reproducibility of this chromatin interactome map, we generated an additional ER- $\alpha$  ChIA-PET library by using a different antibody against ER- $\alpha$ <sup>10</sup>. For this biological replicate (IHH015F), we obtained  $2.0 \times 10^7$  PET sequences (Table 1 and Supplementary Methods). Overall, the two ER- $\alpha$  ChIA-PET libraries were very similar, with many overlapping ER- $\alpha$ BSs and intrachromosomal interactions but few interchromosomal interactions (Table 1 and Supplementary Tables 1 and 2). The ER- $\alpha$ BSs identified in these two libraries showed high reproducibility, especially for highly enriched binding peaks. The 2,513 ER- $\alpha$ BSs with at least 50 PET counts per cluster (high enrichment) overlapped with more than 70% of the ER- $\alpha$ BSs in the replicate ChIA-PET library (Supplementary Table 6). Furthermore, these high-enrichment ER- $\alpha$ BSs intersected well with previously reported ER- $\alpha$  binding maps<sup>9,10</sup> (Fig. 2a and Supplementary Fig. 13). High-enrichment ER- $\alpha$ BSs are therefore more reliable than low-enrichment sites. Many intrachromosomal interaction regions are detected in both replicate libraries. Highly abundant chromatin interactions are mostly reproducible. Of the top 100 most abundant chromatin interactions in IHM001F, 86 were found in IHH015F (more analyses are given in Supplementary Table 7). Furthermore, all interactions previously identified and validated in this study are found in both replicate libraries (Supplementary Table 5). Conversely, none of the putative interchromosomal interactions were reproducible.

Taken together, our results demonstrate that the ChIA-PET method is highly reliable. Furthermore, our data suggest that ER- $\alpha$  functions primarily by means of an intrachromosomal mechanism. Our subsequent analyses therefore focused on intrachromosomal interactions.

Downstream analyses for both ChIA-PET replicate libraries showed similar results; for simplicity, we discuss our results here using IHM001F, but results for IHH015F are given in Supplementary Note 4. We examined how many ER- $\alpha$ BSs are involved in complex and duplex interactions, or in no interactions (Fig. 2b–d). Our analysis



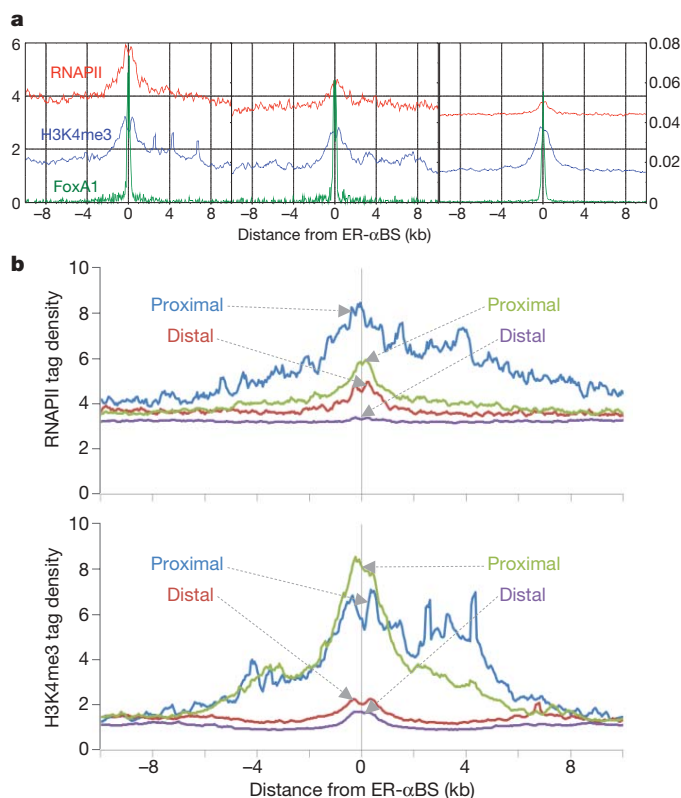
**Figure 2 | ER- $\alpha$ BS reproducibility and association with chromatin interactions.** **a**, Numbers of ER- $\alpha$ BSs identified with different ChIP enrichment cutoffs and reproducibility analyses as measured by overlapping with another ChIA-PET dataset (IHH015F), ChIP-Seq<sup>10</sup> and ChIP-chip<sup>7</sup> data. Left: low-enrichment ER- $\alpha$ BSs, 11,955 (5–49 PETs per site). Right: high-enrichment ER- $\alpha$ BSs, 2,513 (50 or more PETs per site). **b–d**, Examples of ER- $\alpha$ BSs involved in complex interactions (**b**), duplex interactions (**c**) and no interactions (**d**) (singleton inter-ligation PETs only (right) or no inter-ligation PETs (left)). **e**, ER- $\alpha$ BS distribution in different categories of interactions as exemplified in **b–d**. Left: low-enrichment ER- $\alpha$ BSs. Right: high-enrichment ER- $\alpha$ BSs.



showed that high-enrichment ER- $\alpha$ BSs are much more frequently involved in interactions (53%) than are low-enrichment ER- $\alpha$ BSs (only 9%) (Fig. 2e and Supplementary Fig. 13), suggesting that high-confidence and strong ER- $\alpha$ BSs are more likely than weaker ER- $\alpha$ BSs to be involved in chromatin interactions. To gain a better understanding of ER- $\alpha$ BSs with respect to ER- $\alpha$  target genes, we analysed how many ER- $\alpha$ BSs are proximal or distal to gene promoters, based on a cutoff of 5 kb from transcription start sites (TSSs) of University of California Santa Cruz (UCSC) gene database<sup>26</sup>. Of 2,342 ER- $\alpha$ BSs involved in chromatin interactions, 387 (17%) were proximal and 1,955 (83%) were distal to TSSs (Supplementary Fig. 14). We also observed the same ratio for non-interacting ER- $\alpha$ BSs: 2,043 (17%) were proximal and 10,175 (83%) were distal. Most ER- $\alpha$ BSs are therefore distal to gene TSSs, which is in agreement with previous studies<sup>7,8,10</sup>.

### Chromatin interaction and transcriptional regulation

To investigate the functions of ER- $\alpha$ BSs and ER- $\alpha$ -bound chromatin interactions in transcription activation, we generated genome-wide maps of H3K4me3 and RNAPII ChIP-Seq data from MCF-7 cells under oestrogen induction (Supplementary Methods). H3K4me3 is a histone modification that specifically marks active promoters<sup>27</sup>, and the presence of RNAPII is strong evidence for genes that are actively transcribing<sup>28</sup>. We also analysed previously reported FoxA1 ChIP-Chip data<sup>9</sup>, because FoxA1 is an important cofactor of ER- $\alpha$ <sup>6,9</sup>. Generally, H3K4me3, RNAPII and FoxA1 marks showed enrichment around ER- $\alpha$ BSs in our analyses (Fig. 3a). When we compared interacting ER- $\alpha$ BSs with non-interacting ER- $\alpha$ BSs, we found a significant enrichment gradient of RNAPII and FoxA1 binding around ER- $\alpha$ BSs:



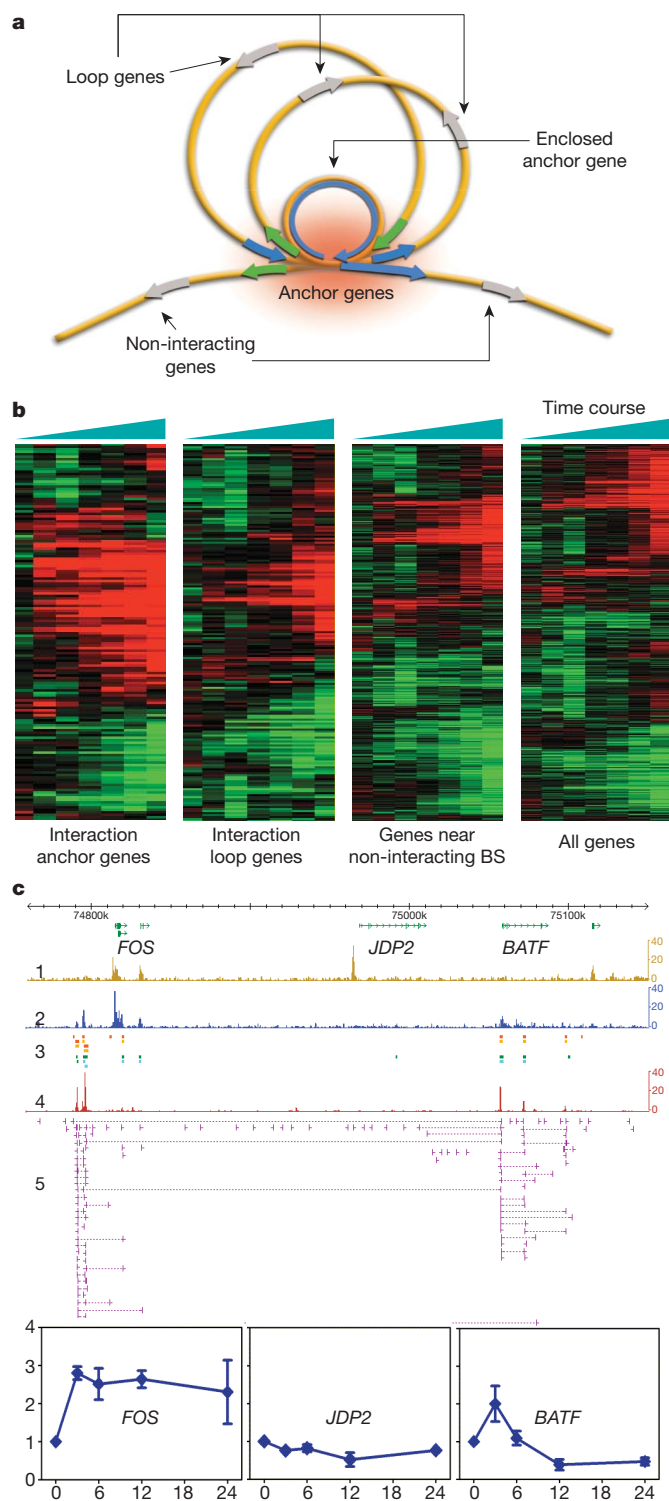
**Figure 3 | Association of ER- $\alpha$ -bound chromatin interactions with functional marks.** **a**, Association of ER- $\alpha$ BSs in complex-interaction (left), duplex-interaction (middle) and non-interacting (right) categories with RNAPII, H3K4me3 and FoxA1 functional marks. The left y axis is relative tag density for H3K4me3 and RNAPII ChIP-seq data; the right y axis is ChIP-chip intensity for FoxA1 ChIP-chip data. **b**, Association of proximal and distal interacting (left) and non-interacting (right) ER- $\alpha$ BSs with RNAPII (top) and H3K4me3 (bottom) functional marks.

most association was with ER- $\alpha$ BSs involved in complex interactions, followed by duplex interactions, and lastly no interactions (Fig. 3a and Supplementary Fig. 15a; significance tests are shown in Supplementary Note 5).

Next we examined the H3K4me3, RNAPII and FoxA1 marks with respect to ER- $\alpha$ BSs proximal or distal to gene promoters and their involvement in chromatin interactions. Proximal ER- $\alpha$ BSs, whether involved in interactions or not, were highly enriched in H3K4me3, but this was not true of distal ER- $\alpha$ BSs, which was expected because H3K4me3 is a known mark for promoter regions (Fig. 3b and Supplementary Fig. 15b; significance tests are shown in Supplementary Note 5). Proximal ER- $\alpha$ BSs were also highly enriched with RNAPII marks, but the enrichment for both proximal and distal ER- $\alpha$ BSs involved in interactions was significantly higher than that of the proximal and distal ER- $\alpha$ BSs that are not involved in interactions. Intriguingly, although RNAPII showed less enrichment around distal interacting ER- $\alpha$ BSs compared with proximal interacting ER- $\alpha$ BSs, the enrichment was significantly higher than that with distal non-interacting ER- $\alpha$ BSs. Conversely, FoxA1 binding was more enriched around distal ER- $\alpha$ BSs than around proximal ER- $\alpha$ BSs, and most enriched around interacting distal ER- $\alpha$ BSs (Supplementary Fig. 15c), and differences were statistically significant (significance tests are shown in Supplementary Note 5). This indicates that RNAPII and FoxA1, but not H3K4me3, predict interactions at distal ER- $\alpha$ BSs, and suggests that RNAPII and FoxA1 participate in tethering chromatin interactions. Whereas RNAPII is strongly associated with ER- $\alpha$ BSs for transcription activation, FoxA1 is more directly correlated with the regulatory function of ER- $\alpha$  at distal ER- $\alpha$ BSs. At least six interacting ER- $\alpha$ BSs bracket the *FOXA1* gene, signifying ER- $\alpha$ -mediated chromatin interactions may regulate *FOXA1* (Fig. 2b), further supporting the hypothesis that FoxA1 and ER- $\alpha$  may regulate each other<sup>29</sup>.

Subsequently, we examined the 689 ER- $\alpha$ -bound chromatin interaction regions with regard to looping structure and gene transcription. We envisage that multiple ER- $\alpha$ BSs may function as 'anchor' regions forming chromatin looping structures in three-dimensional space (Fig. 4a). Genes close to interaction anchors are considered to be 'anchor genes', and genes in the interaction loop regions and faraway from anchors to be 'loop genes'. We annotated the interaction regions in relation to UCSC gene database transcripts<sup>26</sup> (a gene may have multiple transcripts; here we report transcript numbers, but gene numbers are given in Supplementary Note 6). A gene was considered to be associated with a chromatin interaction region if the TSS of a gene was within 20 kb of the interaction boundaries (Supplementary Fig. 14), a parameter that includes many known and validated ER- $\alpha$  target genes. Most interaction regions (393 out of 689 = 57%) were associated with 'anchor genes' (TSS to interaction anchor within 20 kb). Altogether, 1,575 'anchor genes' and 3,767 'loop genes' (TSS more than 20 kb away from interaction anchors) were assigned to interaction regions (Supplementary Tables 3 and 8). Using the same distance parameter (20 kb), we assigned 11,790 genes to 12,126 stand-alone ER- $\alpha$ BSs not involved in interactions (Supplementary Note 6).

Within interaction regions with at least one anchor gene there are 1,073 distal ER- $\alpha$ BSs and 387 proximal ER- $\alpha$ BSs (less than 5 kb to TSS), and all distal ER- $\alpha$ BSs (5' or 3' to the gene promoter) are looped to anchor genes through connections with proximal ER- $\alpha$ BSs. Many interaction regions include multiple genes, such as the keratin gene cluster (Fig. 1c) and the *NR2F2* locus (Fig. 1d), whereas others include only single genes, such as *SLAH2* (Fig. 1b). Distal ER- $\alpha$ BSs are stronger than proximal ER- $\alpha$ BSs; this is the inverse of RNAPII marks, which are stronger at gene promoters than at distal regions (Supplementary Fig. 16; examples are shown in Fig. 1 and Supplementary Fig. 17). These observations suggest that direct ER- $\alpha$  binding might be initiated primarily at one or multiple distal sites, which then subsequently recruit other binding sites as anchors to form an interaction complex to ultimately engage the transcriptional machinery at gene promoters.



**Figure 4 | Proposed ER- $\alpha$ -bound chromatin interaction and transcription regulation mechanism.** **a**, Distal ER- $\alpha$ BSs interact with proximal sites, forming chromatin loops. Anchor genes (green and blue) are close to interaction anchors with concentrated active transcriptional machinery (red shading). Other genes far from interaction centres (grey) are less active. **b**, Expression microarray data (oestrogen induction from 0 to 48 h; red denotes activation, and green repression) for interaction anchor genes, loop genes and genes near non-interacting ER- $\alpha$ BSs, with all other UCSC genes<sup>26</sup>. 'All genes' denotes background. **c**, ChIA-PET interactions data at the *FOS*/*JDP2*/*BATF* loci. Transcription activities are shown by H3K4me3/RNAPII ChIP-Seq and RT-qPCR analysis (bottom panels, the x axis is oestrogen induction time points from 0 to 24 h; the y axis is relative expression (fold increase over ET control)). Error bars show s.e.m.

In addition, we found 296 interaction regions with no associated anchor genes. Although 41 regions contain loop genes, the remaining 255 have no associated UCSC genes assigned to them. Although some interaction regions could be noise or non-functional, some interactions are near gene promoters just outside the 20 kb cutoff, and further sequencing might extend the interaction data to the promoters. The presence of H3K4me3, RNAPII marks and RT-qPCR data at the interaction anchor sites suggests that some interactions could be involved in regulating yet-to-be identified transcripts, such as computationally predicted genes and non-coding RNA species (Supplementary Fig. 18). Alternatively, such interactions could be associated with maintaining chromatin structures or other unknown functions.

To understand whether genes associated with ER- $\alpha$ -bound interactions are regulated by oestrogen, we analysed expression profiles of several interaction-associated genes by RT-qPCR over a time course of oestrogen induction (Supplementary Methods). All anchor genes examined are upregulated by oestrogen induction (Supplementary Fig. 8). We extended our analysis to all interaction-associated genes with the use of whole-genome gene expression microarrays (Fig. 4b). Most 'anchor genes' are upregulated (60%), particularly at early time points, in comparison with 'loop genes' (48%), indicating that 'anchor genes' are significantly associated with gene upregulation (two-tailed  $P = 1.25 \times 10^{-16}$ ; Fig. 4c, Supplementary Note 7, Supplementary Table 9 and Supplementary Fig. 19). In addition, RNAPII marks are associated more with 'anchor genes' (39%) than with 'loop genes' (26%) (two-tailed  $P = 10^{-19}$ ). Conversely, genes assigned to ER- $\alpha$ BSs not involved in interactions (on the basis that the gene promoters are within 20 kb of non-interacting ER- $\alpha$ BSs) have very similar expression profiles to that of the background control (all UCSC genes not associated with interactions), indicating that genes associated with non-interacting ER- $\alpha$ BSs are less activated than genes associated with interaction ER- $\alpha$ BSs (significance tests are shown in Supplementary Note 7). Hence, some stand-alone ER- $\alpha$ BSs could be noise, whereas others could involve non-looping mechanisms such as the recruitment of secondary coactivators for downstream functions<sup>6</sup>.

Within the anchor gene category, many (495 out of 1,575 = 31%) gene entries have 5' and 3' ends within interaction boundaries. Such entries, called 'enclosed anchor genes', frequently occupy the entirety of short interaction loops, engage multiple anchor sites around or within the gene, tend to have intense RNAPII marks covering the entire gene (examples are shown in Fig. 2b, c and Supplementary Fig. 20) and are preferentially associated with RNAPII marks and gene upregulation as indicated by expression microarrays (Supplementary Note 7 and Supplementary Table 9).

Taken together, our data show an association between chromatin interactions and gene transcriptional activation: enclosed anchor genes are closely correlated with upregulation as measured by gene expression microarray data and RNAPII ChIP-Seq peaks; less closely correlated are non-enclosed anchor genes, loop genes even less so, and genes not associated with interactions are much less so. These results suggest that gene-centric interaction structures may enclose a compartment for concentrating ER- $\alpha$  and transcription-related proteins at target genes.

ER- $\alpha$ -bound interactions may coordinate transcriptional regulation for multiple genes involved in the same functional pathways. At the keratin gene cluster interaction loci (Fig. 1c), enclosed anchor genes such as *KRT8* and *KRT18* are actively transcribing, as demonstrated by RNAPII and H3K4me3 marks, whereas the loop genes such as *KRT72* and *KRT75*, which are mainly keratins expressed in hair cells that do not have a function in mammary epithelial cells such as MCF-7, are mostly inactive (Supplementary Note 8). Another example is the complex interaction that encompasses the three genes *FOS*, *JDP2* and *BATF* (Fig. 4c), which encode the dimerization partners of *JUN* to form the AP-1 transcription factors. AP-1 is important in regulating oestrogen-receptor-dependent transcription by functioning either as a DNA tethering partner or as an ER- $\alpha$  cofactor<sup>30</sup>. In



this complex interaction, *FOS* and *BATF* are enclosed anchor genes and are upregulated as shown by RNAPII marks and RT-qPCR, whereas *JDP2* is a loop gene and is downregulated as shown by RT-qPCR and decreased RNAPII occupancy. The promoter of *JDP2* is marked by H3K4me3, a common feature found in many loop genes (Supplementary Table 9). *JDP2* and other loop genes could be 'poised' for activation if they were to escape from the interaction loop. Long-range transcriptional regulation by ER- $\alpha$  may therefore be a fine-tuning mechanism that evolved to regulate specific sets of related genes differentially.

To determine functionally whether some ER- $\alpha$ -associated interaction regions are dependent on ER- $\alpha$ , we used short interfering RNA (siRNA) to knock down the level of ER- $\alpha$  protein in MCF-7 cells (Supplementary Methods) and then measured whether the interactions and gene transcription were affected. ER- $\alpha$ -specific siRNA (siER- $\alpha$ ) efficiently decreased the amount of ER- $\alpha$  protein and effectively

abolished the interactions as demonstrated by a set of 3C assays at the *GREB1* locus (Fig. 5). Furthermore, siER- $\alpha$  blocked *GREB1* transcription as determined by RT-qPCR. Similar results were also previously shown at the *TFF1* site<sup>31</sup>. Together, these data suggest that at least some of the regulatory long-range chromatin interactions identified by ER- $\alpha$  ChIA-PET data are mediated by ER- $\alpha$ .

## Discussion

We demonstrated the ChIA-PET mapping strategy is an unbiased whole-genome approach for the *de novo* analysis of chromatin interactions, and hence is a major technological advance in our ability to study higher-order organization of chromosomal structures and functions. The ChIA-PET interaction data greatly increase the accuracy of assigning distal TFBSs to target genes, and globally addresses the three-dimensional chromatin interaction mechanism by which distal TFBSs regulate transcription. We postulate the following primary mechanism for ER- $\alpha$  function: ER- $\alpha$  protein dimers are recruited to multiple and primarily distal ER- $\alpha$ BSs, which interact with one another and possibly with other factors such as FoxA1 and RNAPII to form chromatin looping structures around target genes; such topological architectures may partition individual genes into subcompartments of nuclear space such as interaction-anchor-associated genes and interaction-loop-associated genes for differential transcriptional activation or repression. We further speculate that tightly enclosed chromatin interaction centres could help achieve and maintain high local concentration of transcription components for efficient cycling of transcriptional machinery on target gene templates (a summary of results is given in Supplementary Information, and more discussion in Supplementary Note 9).

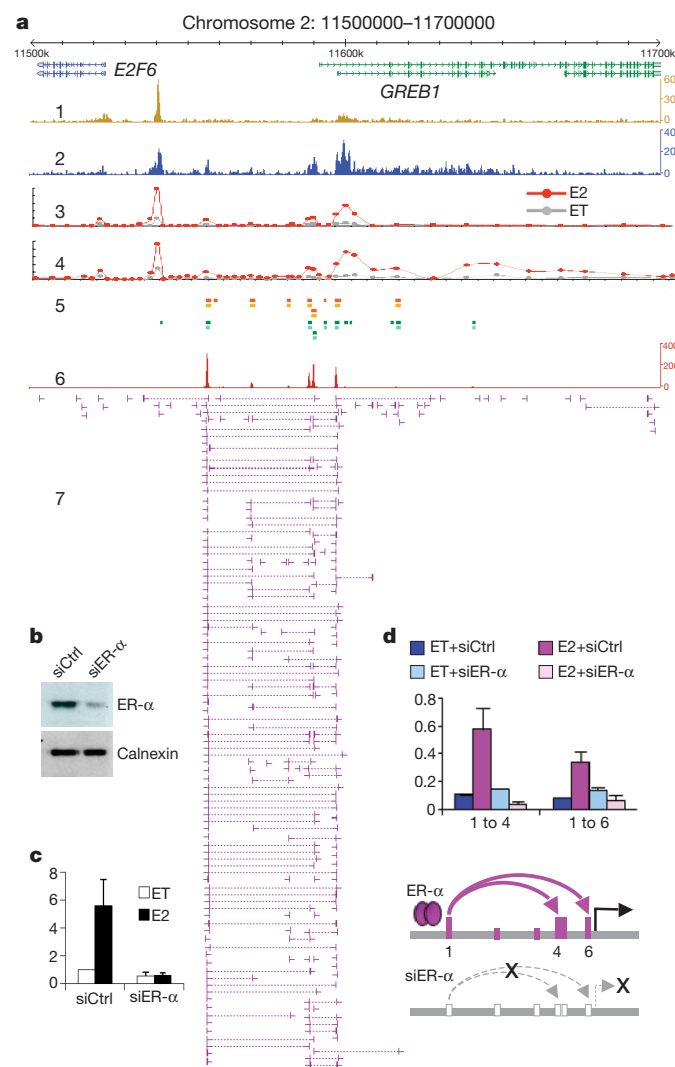
We expect that this global chromatin interactome map and the ChIA-PET assay will be a valuable starting point for future studies of the three-dimensional architecture of transcription biology in whole-genome contexts.

## METHODS SUMMARY

MCF-7 cells grown in hormone-depleted medium were treated with 17 $\beta$ -oestradiol ('oestrogen', E2) for 45 min before being crosslinked with 1% formaldehyde for 10 min. ChIA-PET libraries were constructed by first performing ChIP with HC-20 antibody (Santa Cruz) or Mab-NRF3A6-050 antibody (Diagenode)<sup>10</sup> against ER- $\alpha$ . DNA fragments in ChIP complexes were then ligated to biotinylated half-linkers (linker ligation) containing flanking MmeI restriction sites. The complexes were further ligated under dilute conditions (proximity ligation). PETs were extracted from the ligation products by digestion with MmeI. Released biotinylated PETs were purified by streptavidin-coated magnetic beads, ligated to adaptors, and amplified by PCR. Gel-purified amplicons of PET templates were sequenced by Roche/454 and/or Illumina paired-end sequencing. PET sequences were mapped to the human reference genome (hg18). Binding sites and interactions were identified by using a readout of overlap PETs. To correct for ChIP enrichment bias, we formulated a statistical analysis framework to calculate the probability of the formation of inter-ligation PETs between two regions if ligations between DNA fragments occur by chance. Interactions were further collapsed into complex interactions if they shared interaction anchors. UCSC genes were assigned to interaction regions if they were within 20 kb of interaction regions. To characterize ER- $\alpha$ -bound interactions and associated genes functionally, we conducted gene expression microarray experiments in a time course with and without E2 treatment, and generated genome-wide maps of H3K4me3 (ab8580; Abcam) and RNAPII (serine-5 phosphorylation antibody, ab5131; Abcam) ChIP-Seq data by using Illumina GA single-read sequencing. Interaction-associated genes were annotated with expression microarray data and RNAPII and H3K4me3 ChIP-Seq peaks. Validation experiments included ChIP-qPCR, 3C, ChIP-3C, 4C, FISH and RT-qPCR. For siRNA studies, ER- $\alpha$  ON-TARGETplus SMARTpool siRNA (Dharmacon) was transfected into MCF-7 cells with Lipofectamine 2000 (Invitrogen). Sequences used in experiments are listed in Supplementary Table 10.

Received 25 July; accepted 11 September 2009.

- Fraser, P. Transcriptional control thrown for a loop. *Curr. Opin. Genet. Dev.* **16**, 490–495 (2006).
- Collas, P. & Dahl, J. A. Chop it, ChIP it, check it: the current status of chromatin immunoprecipitation. *Front. Biosci.* **13**, 929–943 (2008).



**Figure 5 | ER- $\alpha$ -bound chromatin interactions are required for transcription activation.** **a**, Genome browser at the *GREB1* locus showing data tracks: 1 and 2, H3K4me3 and RNAPII ChIP-Seq; 3 and 4, RNAPII ChIP-qPCR scans using different RNAPII antibodies under oestrogen induction (E2, red) and ethanol control (ET, grey); 5, ER- $\alpha$  (orange) and FoxA1 (green) ChIP-chip<sup>7</sup>; 6, ChIA-PET density; 7, interaction data. **b–d**, siRNA knockdown experiments. MCF-7 cells were transfected with siRNA against ER- $\alpha$  (siER- $\alpha$ ) or control (siCtrl), and then analysed by western blot analysis with ER- $\alpha$  and calnexin (control) antibodies (**b**); RT-qPCR on *GREB1* expression (**c**); and 3C assays at *GREB1* (**d**): siER- $\alpha$  knockdown abolishes chromatin interactions and turns off transcription. Error bars show s.e.m. The y axes are relative expression (**c**) and normalized interaction (**d**).

3. Wei, C. L. *et al.* A global map of p53 transcription-factor binding sites in the human genome. *Cell* **124**, 207–219 (2006).
4. Wold, B. & Myers, R. M. Sequence census methods for functional genomics. *Nature Methods* **5**, 19–21 (2008).
5. Massie, C. E. & Mills, I. G. ChIPping away at gene regulation. *EMBO Rep.* **9**, 337–343 (2008).
6. Carroll, J. S. *et al.* Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**, 33–43 (2005).
7. Carroll, J. S. *et al.* Genome-wide analysis of estrogen receptor binding sites. *Nature Genet.* **38**, 1289–1297 (2006).
8. Lin, C. Y. *et al.* Whole-genome cartography of estrogen receptor  $\alpha$  binding sites. *PLoS Genet.* **3**, e87 (2007).
9. Lupien, M. *et al.* FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**, 958–970 (2008).
10. Welboren, W. J. *et al.* ChIP-Seq of ER $\alpha$  and RNA polymerase II defines genes differentially responding to ligands. *EMBO J.* **28**, 1418–1428 (2009).
11. West, A. G. & Fraser, P. Remote control of gene transcription. *Hum. Mol. Genet.* **14** (Review Issue 1) R101–R111 (2005).
12. Woodcock, C. L. Chromatin architecture. *Curr. Opin. Struct. Biol.* **16**, 213–220 (2006).
13. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295**, 1306–1311 (2002).
14. Hagege, H. *et al.* Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nature Protocols* **2**, 1722–1733 (2007).
15. Horike, S., Cai, S., Miyano, M., Cheng, J. F. & Kohwi-Shigematsu, T. Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nature Genet.* **37**, 31–40 (2005).
16. Cai, S., Lee, C. C. & Kohwi-Shigematsu, T. SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nature Genet.* **38**, 1278–1288 (2006).
17. Zhao, Z. *et al.* Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nature Genet.* **38**, 1341–1347 (2006).
18. Ling, J. Q. *et al.* CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* **312**, 269–272 (2006).
19. Simonis, M. *et al.* Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nature Genet.* **38**, 1348–1354 (2006).
20. Wurtele, H. & Chartrand, P. Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. *Chromosome Res.* **14**, 477–495 (2006).
21. Dostie, J. *et al.* Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16**, 1299–1309 (2006).
22. Tiwari, V. K., Cope, L., McGarvey, K. M., Ohm, J. E. & Baylin, S. B. A novel 6C assay uncovers Polycomb-mediated higher order chromatin conformations. *Genome Res.* **18**, 1171–1179 (2008).
23. Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F. & Fraser, P. Long-range chromatin regulatory interactions *in vivo*. *Nature Genet.* **32**, 623–626 (2002).
24. Osborne, C. S. *et al.* Active genes dynamically colocalize to shared sites of ongoing transcription. *Nature Genet.* **36**, 1065–1071 (2004).
25. Simonis, M., Kooren, J. & de Laat, W. An evaluation of 3C-based methods to capture DNA interactions. *Nature Methods* **4**, 895–901 (2007).
26. Hsu, F. *et al.* The UCSC Known Genes. *Bioinformatics* **22**, 1036–1046 (2006).
27. Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
28. Phatnani, H. P. & Greenleaf, A. L. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* **20**, 2922–2936 (2006).
29. Laganier, J. *et al.* Location analysis of estrogen receptor  $\alpha$  target promoters reveals that FOXA1 defines a domain of the estrogen response. *Proc. Natl Acad. Sci. USA* **102**, 11651–11656 (2005).
30. Kushner, P. J. *et al.* Estrogen receptor pathways to AP-1. *J. Steroid Biochem. Mol. Biol.* **74**, 311–317 (2000).
31. Pan, Y. F. *et al.* Regulation of estrogen receptor-mediated long-range transcription via evolutionarily conserved distal response elements. *J. Biol. Chem.* **283**, 32977–32988 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature). A summary figure is also included.

**Acknowledgements** The authors acknowledge the Genome Technology and Biology Group at the Genome Institute of Singapore for technical support; A. Shahab, C. C. Seng and F. H. Mulawadi for computing support; S. Luo and G. Schroth for Illumina sequencing support; and W. de Laat, B. Ren and X. S. Liu for advice. M.J.F., P.Y.H.H., Y.H., P.Y.T. and Y.K.L. are supported by A\*STAR Scholarships. M.J.F. is supported by a L'Oreal-UNESCO For Women In Science National Fellowship. Y.R. and C.L.W. are supported by A\*STAR of Singapore and NIH ENCODE grants (R01 HG004456-01, R01HG003521-01 and part of 1U54HG004557-01).

**Author Contributions** M.J.F. and Y.R. conceptualized the ChIA-PET strategy. M.J.F., E.C. and Y.R. designed the experiments. M.J.F., M.H.L., Y.F.P., J.L., A.H., P.H.M., E.G.Y.C., P.Y.Y.H., W.-J.W., Y.H., Y.L., P.Y.T., P.Y.C., K.D.S.A.W., B.Z., K.S.L., S.C.L., J.S.Y., R.J., K.V.D., J.S.T., Y.K.L., T.H., H.G.S., X.R. and V.C.-R. performed experiments. M.J.F., X.H., Y.B.M., Y.L.O., S.V., H.-S.O., P.N.A., V.B.V., Y.K.L., R.K.M.K., G.B., H.G.S., X.R., V.C.-R., W.-K.S., C.-L.W., E.C. and Y.R. analysed experimental data. E.T.L., E.C. and C.-L.W. commented on the manuscript drafts; M.J.F. and Y.R. wrote the paper.

**Author Information** The data in the ChIA-PET libraries have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE18046. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to Y.R. ([ruanyj@gis.a-star.edu.sg](mailto:ruanyj@gis.a-star.edu.sg)) or E.C. ([cheungcwe@gis.a-star.edu.sg](mailto:cheungcwe@gis.a-star.edu.sg)).



# Combinatorial binding predicts spatio-temporal *cis*-regulatory activity

Robert P. Zinzen<sup>1\*</sup>, Charles Girardot<sup>1\*</sup>, Julien Gagneur<sup>1\*</sup>, Martina Braun<sup>1</sup> & Eileen E. M. Furlong<sup>1</sup>

**Development requires the establishment of precise patterns of gene expression, which are primarily controlled by transcription factors binding to *cis*-regulatory modules. Although transcription factor occupancy can now be identified at genome-wide scales, decoding this regulatory landscape remains a daunting challenge. Here we used a novel approach to predict spatio-temporal *cis*-regulatory activity based only on *in vivo* transcription factor binding and enhancer activity data. We generated a high-resolution atlas of *cis*-regulatory modules describing their temporal and combinatorial occupancy during *Drosophila* mesoderm development. The binding profiles of *cis*-regulatory modules with characterized expression were used to train support vector machines to predict five spatio-temporal expression patterns. *In vivo* transgenic reporter assays demonstrate the high accuracy of these predictions and reveal an unanticipated plasticity in transcription factor binding leading to similar expression. This data-driven approach does not require previous knowledge of transcription factor sequence affinity, function or expression, making it widely applicable.**

Gene expression states are established through the integration of signalling and transcriptional networks converging on enhancer elements, also known as *cis*-regulatory modules (CRMs)<sup>1,2</sup>. CRMs integrate the input of multiple transcription factors, leading to a specific spatio-temporal output of expression<sup>3</sup>, and are therefore central to understanding gene regulation and metazoan development. Although there has been considerable progress in decoding individual CRM activities<sup>4–15</sup>, it is difficult to scale these approaches to decipher entire transcriptional networks at a genomic level. Understanding global *cis*-regulatory networks requires a detailed knowledge of the location of all developmental CRMs, a comprehensive map of their combinatorial and temporal binding profiles and the ability to predict their spatio-temporal activity.

Chromatin immunoprecipitation followed by either microarray analysis (ChIP-on-chip) or sequencing enables the identification of active CRMs in an unbiased, genome-wide and systematic manner. For example, ChIP experiments with tissue-specific factors in the context of the entire embryo<sup>16–20</sup>, or against factors in isolated tissues<sup>21,22</sup>, have proven to be very accurate methods for identifying tissue-specific CRMs. Many of these studies focus on the occupancy of a single factor or a snapshot in time<sup>16,19,23,24</sup>; however, transcription factors rarely act in isolation, but rather bind to CRMs in a combinatorial and dynamic manner to regulate specific expression patterns. Thus, a global view of how the combinatorial binding of transcription factors translates into specific spatio-temporal expression patterns is still lacking.

Given the accuracy of ChIP approaches<sup>16–22</sup>, the future challenge is to move beyond predicting the location of CRMs to rather predict their spatio-temporal activity. A number of sequence-based models have been applied to predict spatio-temporal activity of enhancers during development<sup>14,25–27</sup>. These methods are very accurate when tailored to individual CRMs<sup>14,26</sup> or even small numbers of regulatory modules<sup>27</sup>; however, they rely on detailed knowledge of the system, including estimates of transcription factor concentrations, their affinity for various sequence motifs, and cooperativity or competition between transcription factors<sup>14,26,27</sup>—data that are currently only available for a handful of CRMs. Here we present a complementary,

data-driven approach to predict spatio-temporal CRM activity using only transcription factor binding and *in vivo* activity data as input to a machine-learning algorithm. We applied this approach to the transcriptional network governing the specification of the *Drosophila* mesoderm into different muscle primordia.

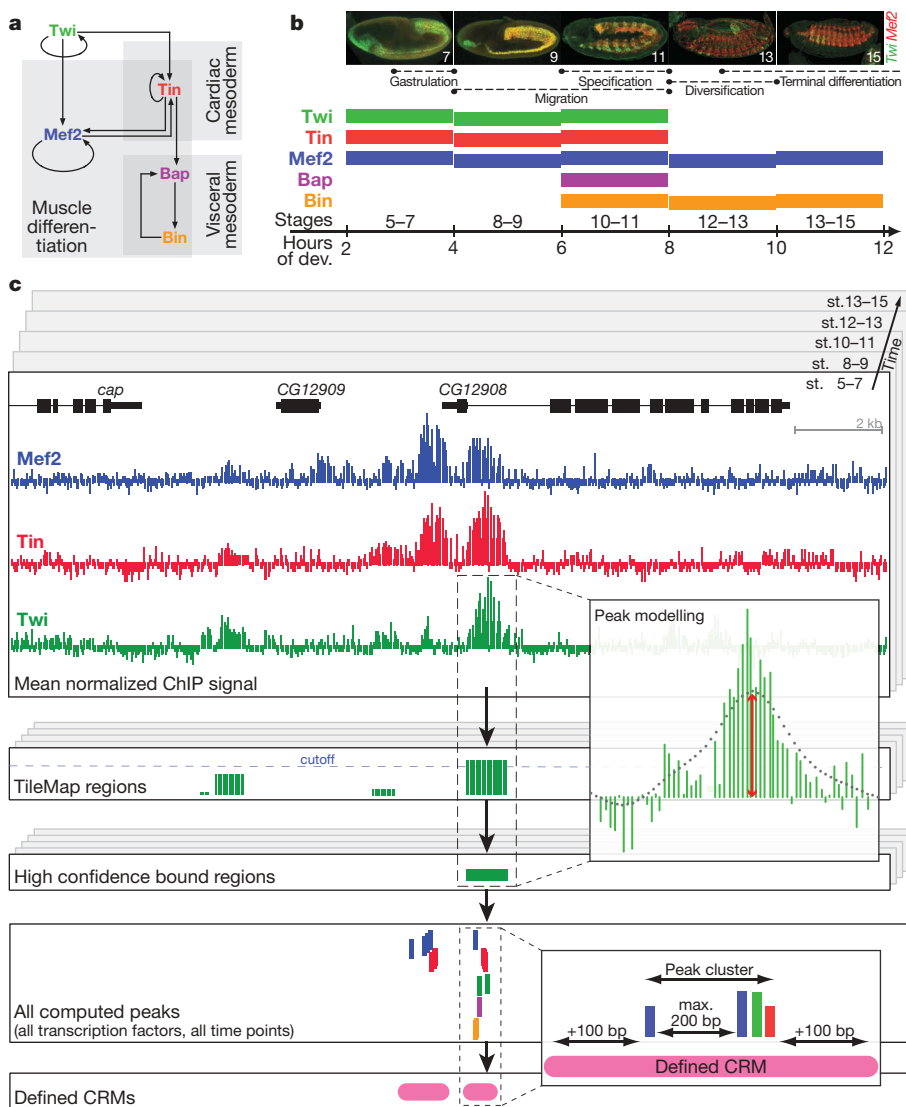
## A high-resolution binding atlas of mesoderm development

The subdivision of the mesoderm requires the successive activation of transcription factors whose activities result in the specification of cardiac mesoderm (heart muscle), somatic muscle (analogous to vertebrate skeletal muscle) and visceral muscle (gut muscle) primordia. The basic helix–loop–helix factor Twist (Twi) acts as a ‘master’ regulator of mesoderm development (Fig. 1a)<sup>28</sup>. Twi directly regulates the expression of both Tinman (Tin)<sup>29</sup>, which is essential for dorsal mesoderm specification<sup>30</sup>, and Myocyte enhancing factor 2 (Mef2), which initiates muscle differentiation<sup>31,32</sup>. Tin in turn regulates Bagpipe (Bap) expression<sup>13</sup>, which acts together with Biniou (Bin) to specify the visceral muscle<sup>19,33</sup>. Although many other transcription factors act to refine further the specification and subsequent differentiation of specific muscle types, these five key transcription factors (Fig. 1a) act as the high-level regulators of mesoderm development<sup>34,35</sup> and therefore serve as a good entry point to generate a global atlas of mesodermal CRMs. We previously examined the binding profiles of these transcription factors individually<sup>16,18,19,36</sup> but at a resolution too low to map combinatorial binding precisely or to model CRM activity. We therefore initiated this study by generating a high-resolution, genome-wide map of transcription factor occupancy for these five factors during multiple stages of mesoderm development.

ChIP-on-chip was performed on each transcription factor at consecutive time points spanning the majority of stages when each transcription factor is expressed, resulting in binding data for 15 developmental conditions (Fig. 1b). To minimize false positives owing to potential off-target effects, two independent antibodies for each transcription factor were used. The immunoprecipitated material was hybridized to genome-wide tiling arrays and enriched regions were identified as clusters of consecutive probes with significant

<sup>1</sup>European Molecular Biology Laboratory, D-69117 Heidelberg, Germany.

\*These authors contributed equally to this work.



**Figure 1 | Generating a high-resolution atlas of mesodermal CRMs.** **a**, Myogenic network of five central transcription factors in mesoderm specification. **b**, Coloured rectangular boxes indicate consecutive 2-h time windows assayed by ChIP-on-chip for each transcription factor. Major events in mesoderm specification are indicated (top) during the developmental stages assayed (bottom). **c**, Data analysis workflow. Top panel shows normalized mean log<sub>2</sub>-ratios of ChIP-on-chip signal per transcription factor (horizontal tracks) and time point (stacked windows). Significantly bound regions are calculated for all transcription factors and time points using TileMap (shown for Twist at 2–4 h (stages 5–7)). Peaks were calculated as extrema (red arrow in ‘peak modelling’ inset) in selected regions. ChIP CRMs were computed based on peak clustering, indicated in pink.

signal (TileMap)<sup>37</sup>. These experiments identified thousands of high-confidence bound regions (Supplementary Table 1), providing a genome-wide map reflecting the location, temporal occupancy and combinatorial binding of mesodermal transcription factors *in vivo*.

### Clustered transcription factor binding defines CRMs

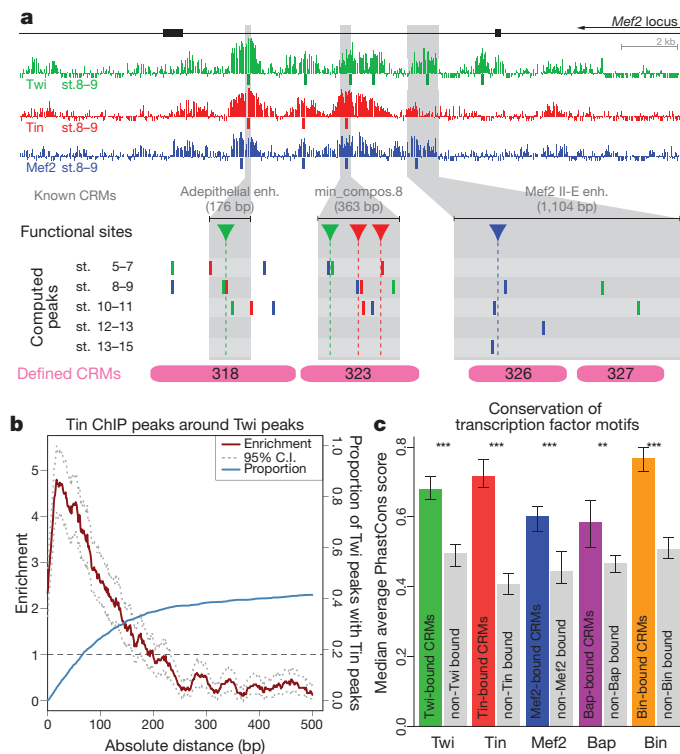
Functional transcription factor binding sites (TFBSs) typically cluster together to form regulatory elements or CRMs; we used this property to search for genomic regions containing clusters of ChIP-binding peaks in close proximity. To determine an appropriate distance for proximity we first assessed the precision of our ChIP data. As TFBSs are generally located beneath ChIP enrichment peaks<sup>20,38,39</sup>, we compared the relative location of computed ChIP peaks<sup>40</sup> (Fig. 1c and Methods) to previously characterized functional TFBSs, and more globally using predicted sites. Mutagenesis analyses in three CRMs regulating *Mef2* expression identified active TFBSs for *Twist*, *Tinman* and *Mef2*<sup>41–43</sup> (Fig. 2a), providing a good test case to assess the ability of ChIP peaks to pinpoint functional sites. Computed ChIP peaks were, on average, within 50 bp of the functional TFBSs (Fig. 2a), demonstrating the high sensitivity and resolution of the data. More globally, the absolute distance of ChIP peaks for a given transcription factor to the closest predicted TFBS is significantly enriched within 100 bp (Supplementary Fig. 1).

We next assessed the relative distance of transcription factor binding events for different factors at the same stage of development. ChIP peaks for different transcription factors are positioned much closer to each other than expected by chance (Fig. 2b and Supplementary

Fig. 2), indicating that these occupied binding sites cluster into regulatory modules. Given the ChIP precision of  $\pm 100$  bp, we defined CRMs as 200-bp windows centred on ChIP peaks, and combined regions where peaks clustered into overlapping windows. The ChIP experiments from all 15 conditions identified 19,522 high-confidence ChIP peaks, which clustered into 8,008 regions; almost half of these (3,713) are multi-peak regions bound by more than one transcription factor, or by one transcription factor at more than one time point (Supplementary Fig. 3). These computed ChIP CRMs cover 2.17 Mb of genomic sequence, representing a 3.3-fold increase in resolution compared to TileMap enriched regions, and have an average length of 270 bp, which is significantly smaller than what is commonly tested in enhancer-reporter assays (Supplementary Fig. 4).

The accuracy of the ChIP-CRM atlas was assessed globally using TFBS conservation as an indicator of regulatory function and more locally using *in vivo* transgenic reporter assays. To assess conservation, we first optimized the position weight matrices (PWMs) for the five transcription factors (Supplementary Fig. 5, 5' and Methods). The refined PWMs were used to search for all instances of these motifs within the *Drosophila* genome. As expected, these motifs are highly enriched within ChIP CRMs, being present within 100 bp of the ChIP peak in ~60–80% of CRMs (Supplementary Fig. 1, blue line). Two methods were used to globally assess TFBS conservation: first, the average PhastCons score<sup>44</sup> across predicted TFBSs within ChIP CRMs for a given transcription factor was compared to that of CRMs not bound by that transcription factor to minimize general sequence biases of CRMs (Fig. 2c)<sup>20</sup>. Second, the percentage of conserved TFBSs derived from pair-wise alignments of





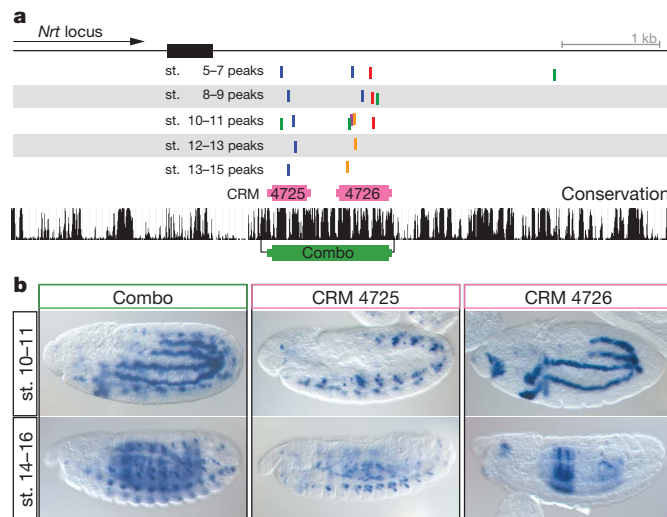
**Figure 2 | ChIP peaks are within 100 bp of transcription factor motifs, which are globally conserved.** **a**, Transcription factor binding in the *Mef2* locus. Three characterized CRMs (grey shading) with functional TFBSs (arrowheads, dotted lines): Twi, green; Tin, red; Mef2, blue. ChIP peaks (vertical bars) are within 31, 26, 101, 52 and 39 bp of verified TFBSs. ChIP CRMs are indicated in pink. **b**, Enrichment (red) and 95% confidence intervals (dotted grey) of Tin peaks at a given base-pair distance from a Twi peak. Blue line shows cumulative production of Twi peaks within Tin peaks. All at stage 8–9. **c**, Conservation of transcription factor motifs: median average PhastCons scores for motifs in bound (coloured bars) and non-bound CRMs for that transcription factor (grey bars). Error bars represent equi-tailed 95% confidence intervals of the median. \*\*\* $P < 0.001$ ; \*\*\* $P < 1.0 \times 10^{-6}$  (one-sided Wilcoxon's rank-sum test).

orthologous CRMs was compared between different *Drosophila* species (Supplementary Fig. 6 and Methods). Both analyses revealed significant TFBS conservation for all five transcription factors, suggesting that the majority of ChIP CRMs are likely to have regulatory function (Fig. 2c and Supplementary Fig. 6).

Notably, 35 out of 36 ChIP CRMs tested during this study are sufficient to function as discrete regulatory modules *in vivo*. These regions were selected based only on their transcription factor binding profiles (see below), irrespective of their evolutionary conservation or the identity of the neighbouring genes. A specific example is the *Neurotactin* locus (Fig. 3a), where several transcription factors bind within the first intron. Taking conservation as a guide, we tested a 1,200-bp region in transgenic reporter assays, which showed specific expression in a somatic muscle subset and in the visceral muscle (Fig. 3b). However, the ChIP-CRM atlas indicated two regulatory modules within this region: a 350-bp CRM containing 6 ChIP peaks (number 4725) and a 502-bp CRM containing 11 ChIP peaks (number 4726) (Fig. 3a). Testing the activities of these regions individually demonstrates that the somatic muscle and visceral muscle expression of the composite CRM are modular and separable, residing in CRM number 4725 and 4726, respectively (Fig. 3b). The high precision of the ChIP-CRM atlas can therefore define discrete regulatory units, facilitating a global analysis of spatio-temporal activity.

### CAD, a CRM activity database

Our motivation for this study was to investigate whether combinatorial transcription factor binding is predictive of CRM activity during



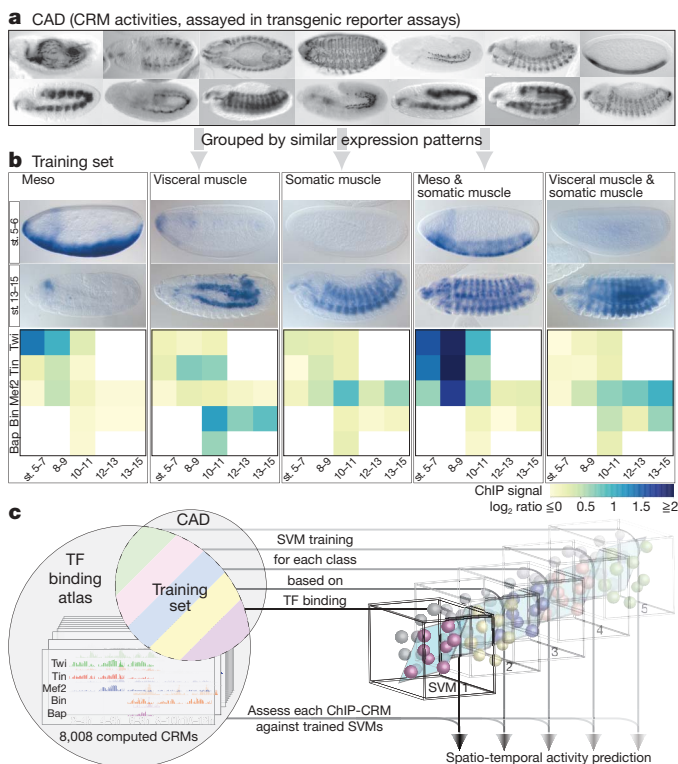
**Figure 3 | ChIP CRMs act as discrete functional units.** **a**, ChIP peaks for five transcription factors at multiple developmental stages within the first intron of *Neurotactin* (*Nrt*; top); peaks are indicated by vertical bars (colours as in Fig. 1b). Two distinct CRMs (numbers 4725 and 4726) were computed in close proximity, indicated in pink above a UCSC conservation track. **b**, *In situ* hybridization of the *lacZ* reporter driven by the 1.2-kb conserved region (green; Combo) and by the individual ChIP CRMs. The somatic muscle and visceral muscle activity (Combo) is modular and separable; CRM 4725 drives somatic muscle and 4726 visceral muscle expression. All embryos are oriented anterior-left, dorsal-up.

tissue differentiation. To facilitate this, we collected a reference data set of enhancers with characterized tissue-specific expression from *in vivo* transgenic reporter assays. The CRM Activity Database (CAD) (Fig. 4a) uses a controlled vocabulary to compile spatio-temporal information about the expression driven by published CRMs, using REDFly 2.0<sup>45</sup>, literature surveys and our own experiments (Supplementary Methods). The annotation was manually reviewed on a CRM-by-CRM basis and overlapping regions were combined, split, or eliminated. This resulted in a collection of 525 largely non-redundant CRMs, 139 of which drive expression in mesoderm and/or muscle (Supplementary Tables 2 and 4).

As the majority of enhancers in CAD were identified in single gene studies, the 139 CRMs provide an independent resource to assess the sensitivity of the ChIP experiments. The ChIP-CRM atlas covers 77% of all known mesodermal-muscle CRMs, demonstrating high sensitivity and providing new insight into their temporal and combinatorial occupancy.

### Using transcription factor binding to predict expression

The intersection of CAD and the ChIP-CRM atlas identified 310 ChIP CRMs for which both their combinatorial binding profile (ChIP atlas) and spatio-temporal expression (CAD) are known. We used these regions, referred to as the training set (Methods and Supplementary Table 8), to train a machine-learning algorithm to predict CRM spatio-temporal expression based on transcription factor binding profiles (Fig. 4). Five exclusive expression classes were defined as groups of CRMs with activity in specific domains, including three single-tissue classes—the early unspecified mesoderm ('Meso') and two muscle derivatives specified later in development ('somatic muscle' or 'visceral muscle')—and two complex classes combining two expression domains—'Meso & somatic muscle' or 'visceral muscle & somatic muscle' (Fig. 4b, Methods and Supplementary Table 3). For each expression class, a support vector machine (SVM) was trained to discriminate between members and non-members of the class given only transcription factor binding data. ChIP-peak heights were used as a quantitative estimate of transcription factor occupancy (Methods). Although differences in peak height may reflect several properties, it provides a biochemical quantification of the enrichment of



**Figure 4 | Predicting CRM activity using a machine-learning approach.**

**a**, CAD contains a largely non-redundant collection of CRMs expression patterns. **b**, CRMs driving expression in the same tissue were grouped into five exclusive classes (Meso, visceral muscle, somatic muscle, Meso & somatic muscle, and visceral muscle & somatic muscle), which were used to train five SVM classifiers. The average transcription factor binding profile is shown as a heat map using peak height as a measure of transcription factor occupancy. **c**, The binding profile of each CRM within the training set was used to train an SVM, which was then run on all ChIP CRMs to predict their activity.

transcription factor binding and yields more accurate predictions than binary transcription factor binding information (Supplementary Fig. 7 and Methods). No other information was supplied to the SVM. The ability to accurately predict expression was assessed by a leave-one-out cross-validation procedure (ROC curves, Methods and Supplementary Fig. 8). SVMs with tuned parameters were trained on the complete training set and applied to all 8,008 ChIP CRMs for expression prediction (Fig. 4c, Methods and Supplementary Table 9).

### Combinatorial binding predicts spatio-temporal activity

The spatio-temporal activity of at least six CRMs per expression class, selected with an SVM specificity score >95%, were tested *in vivo* using transgenic reporter assays. The  $\Phi$ C31 integrase system<sup>46</sup> was used to integrate stably all constructs into a common genomic locus, eliminating positional effects. The expression pattern of the *lacZ* reporter driven by the CRMs was assessed by *in situ* hybridization (ISH) and CRM activity annotations were verified by fluorescent multiplex ISH using tissue-specific markers and by an independent expert (Fig. 5, Supplementary Figs 9a, b–13a, b, Supplementary Table 12a, b and Supplementary Methods).

CRMs were scored as 'correct' if they drove expression in the predicted tissue(s) (and in no other mesodermal tissue) and were therefore co-expressed with the appropriate tissue-specific markers (Supplementary Fig. 9b–13b); 'partial' if active in one of the predicted tissues, but other aspects of the mesodermal prediction did not hold true; or 'fail' if the CRM did not drive expression in any predicted tissue. Expression in non-mesodermal tissues was disregarded, as the SVM was not trained to discriminate against non-mesodermal expression. Using this stringent scoring system, the SVM predictions performed remarkably well (Supplementary Fig. 8): of the 35 individual CRMs

tested, the spatio-temporal predictions of 71.4% (25) were correct, whereas 14.3% (5) worked partially and 14.3% (5) failed.

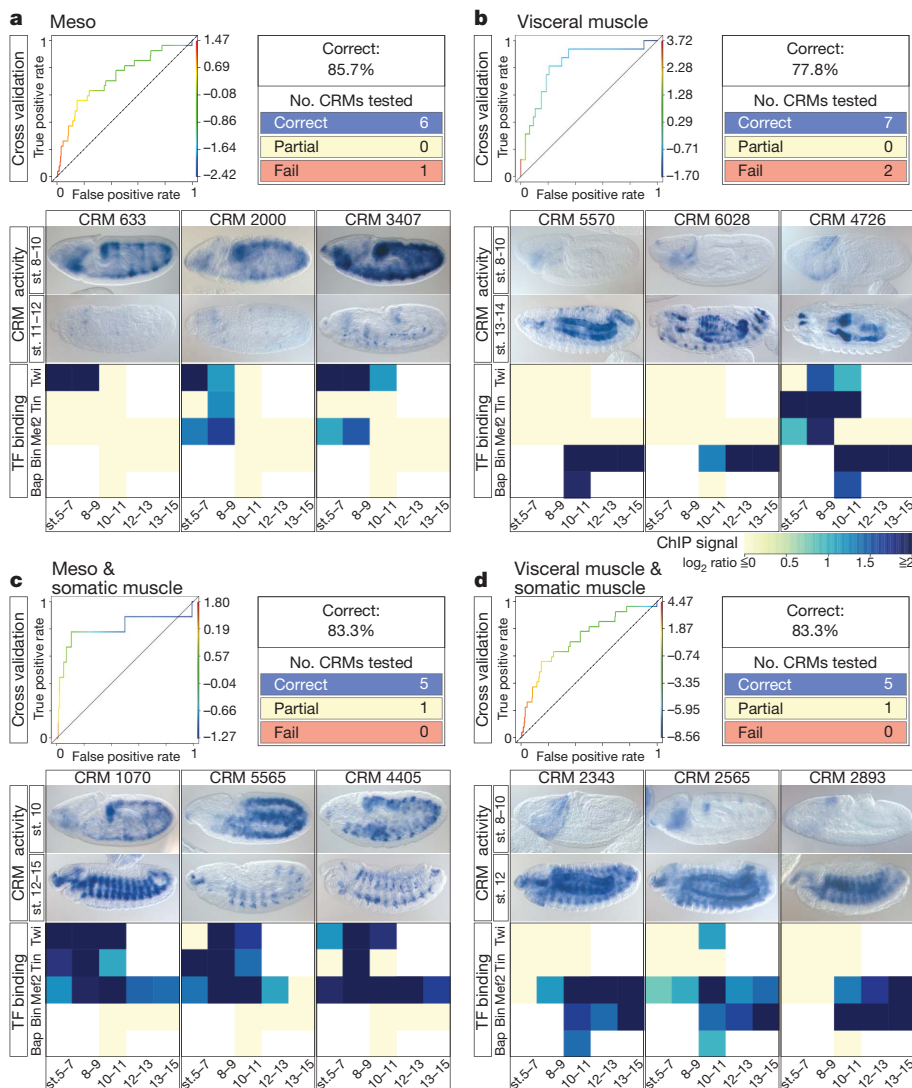
For example, 6 of the 7 CRMs predicted to drive expression in 'Meso' direct expression in the early unspecified mesoderm and not in its derived muscle tissues, even though many transcription factors within the ChIP-CRM atlas are expressed there (Fig. 5a and Supplementary Fig. 9a, b). Similarly, 7 out of 9 'visceral muscle' predictions tested regulate specific expression in visceral muscle and in no other mesodermal tissue (Fig. 5b and Supplementary Fig. 10a, b). Interestingly, a number of these modules drive expression in distinct visceral muscle subsets (Supplementary Fig. 10a, b), indicating input from additional factors to refine their expression.

Importantly, the SVM could also make accurate predictions of more complex spatio-temporal CRM expression involving tissue combinations. In the 'Meso & somatic muscle' class, 5 out of 6 CRMs tested direct expression as predicted in the unspecified mesoderm and somatic muscle, whereas the sixth CRM was partially correct, driving expression early in mesoderm, but not in somatic muscle (Fig. 5c and Supplementary Fig. 12a, b). Similarly, for the tissue combination class 'visceral muscle & somatic muscle', 5 out of 6 CRMs tested direct expression as predicted, whereas the sixth CRM was partially correct, driving expression in visceral muscle, but not in somatic muscle (Fig. 5d and Supplementary Fig. 13a, b).

The predictors of the somatic muscle class were less efficient at recovering training set members in leave-one-out cross validations (ROC plots, Supplementary Fig. 8). This was also reflected in the experimental validation (Supplementary Fig. 11a, b): whereas expression predictions for two out of seven CRMs were correct, predictions for another three CRMs were partially correct, showing expression in somatic muscle and the early mesoderm (CRMs 3775 and 6051) or visceral muscle (CRM 3775 and 6419). The poorer performance of this class probably reflects the inherent complexity of this tissue and the lack of known high-level regulators specific for somatic muscle development<sup>35</sup>. To investigate this further, we compared the binding signature of CRMs within the somatic muscle training set to that of the top SVM predictions (Supplementary Figs 14 and 15). This revealed enrichment in Mef2 binding and depletion in all other transcription factors as the most prominent binding signature for somatic-muscle-associated CRMs. As Mef2 is required for the differentiation of all muscle types<sup>31</sup> it serves as a weak predictor of this class. Nevertheless, even in the absence of binding data for a somatic muscle master regulator(s), the SVM can still recognize some 'somatic muscle signature' as only two CRMs predictions were complete 'fails' (Supplementary Fig. 11a), and the predictions for the 'visceral muscle & somatic muscle' and 'Meso & somatic muscle' classes were very accurate.

Intuitively, CRMs predicted to drive highly similar expression patterns were expected to have very similar transcription factor binding profiles. However, many CRMs with high-ranking predictions in the same tissue have an unexpected diversity in transcription factor occupancy (Fig. 5, Supplementary Fig. 14 and Supplementary Table 9). To demonstrate this, the combinatorial binding profiles of tested CRMs are displayed as 'binding matrices' (Fig. 5), where the intensity of blue represents the quantitative ChIP signal (Supplementary Fig. 14). For all classes, CRMs that drive similar spatio-temporal expression have heterogeneous transcription factor binding profiles in terms of (1) the identity of transcription factors occupying the CRM; (2) the duration of binding; and (3) the intensity of the ChIP signal. For example, three CRMs in the Meso class regulate similar expression in the early mesoderm from stage 5–10 but have quite divergent transcription factor binding profiles (Fig. 5a). This diversity of transcription factor binding, which is also reflected in the training set itself (Supplementary Fig. 15), argues against a stringent combinatorial binding code and probably reflects an unequal contribution of each transcription factor to CRM function. Although some transcription factors may act as the key 'switchers' regulating tissue-specific expression (for example, Twi in early mesoderm), other factors may serve as 'multipliers', fine-tuning the levels of CRM activity.





**Figure 5 | Validation of CRM spatio-temporal predictions *in vivo*.** Results shown for four expression classes: **a**, Meso; **b**, visceral muscle; **c**, Meso & somatic muscle; and **d**, visceral muscle & somatic muscle. **a–d**, Top left: ROC plots of SVM performance in leave-one-out cross validations. Line colour represents SVM score (right y axis). Top right: table summarizing all results of testing SVM predictions. Below: CRM activity in staged embryos, 3 CRM examples per class. Note the lack of somatic muscle or circular visceral muscle staining at stage 11–12 in Meso class (**a**) and lack of mesoderm staining at stage 8–10 in visceral muscle (**b**) and visceral muscle & somatic muscle (**d**) classes. The CRMs' transcription factor (TF) binding profiles are shown as heat maps using log<sub>2</sub> ChIP-peak height (Supplementary Fig. 14). Diverse patterns of transcription factor binding lead to similar CRM expression. ISH data for all CRMs tested are shown in Supplementary Figs 9–13a, b.

## Concluding remarks

Combinatorial binding data are often cited as revealing the regulatory logic of CRMs. However, although binding information is an essential prerequisite, few studies have attempted to bridge the gap from a 'binding code' to the actual regulatory activity. The studies that have modelled developmental CRM activity<sup>14,26,27</sup> used detailed physical models that relied on extensive prior knowledge, including the binding sites within specific CRMs, transcription factor concentrations and transcription factor-DNA affinity estimates. However, this level of information is only available for a very limited number of systems, such as the early patterning of the *Drosophila* embryo. Here we show that combinatorial transcription factor occupancy is sufficient to predict spatio-temporal CRM activity, without the need for prior knowledge of the transcription factors' expression or binding site affinities.

This data-driven approach performed with remarkable accuracy (>70%) given that these transcription factors are expressed in overlapping domains, and that the tissues involved share a developmental history (Fig. 5 and Supplementary Figs 9–13). In the majority of cases, the SVM performed as well or better than a specialist in the field. For example, although the SVM did not have genetic information about the function of these transcription factors, the primary binding signature of CRMs with predicted expression in 'Meso' and 'visceral muscle' was Twi and Bin, respectively, two key transcription factors essential for the development of these tissues (Supplementary Fig. 8). The observed diversity in the occupancy of CRMs regulating similar expression (Fig. 5) questions the generally assumed stringency of regulatory codes. A similar flexibility in regulatory architecture was

observed in enhancers regulating a cohort of 19 co-expressed genes in *Ciona*<sup>47</sup> and thus may represent an inherent property of developmental *cis*-regulatory modules.

The predictive power of this approach will only improve as the activity of more CRMs is tested *in vivo* and more transcription factor occupancy data become available. Given the accumulation of such data in many developmental systems, including higher vertebrates<sup>48,49</sup>, and the robustness of SVMs to accommodate heterogeneous inputs, this method represents a broadly applicable and accurate approach to predict spatio-temporal enhancer expression in complex developmental systems.

## METHODS SUMMARY

ChIP was performed as described previously<sup>50</sup> and hybridized to whole-genome Affymetrix tiling arrays. CRMs were defined as neighbouring clusters of high-confidence transcription factor binding peaks. SVMs were trained with transcription factor binding profiles for five exclusive classes of CRMs with defined activity. SVM predictions were tested using *in vivo* transgenic reporter assays by *in situ* hybridization.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 21 July; accepted 22 September 2009.

- Levine, M. & Davidson, E. H. Gene regulatory networks for development. *Proc. Natl. Acad. Sci. USA* 102, 4936–4942 (2005).
- Ochoa-Espinosa, A. & Small, S. Developmental mechanisms and *cis*-regulatory codes. *Curr. Opin. Genet. Dev.* 16, 165–170 (2006).

3. Arnosti, D. N. & Kulkarni, M. M. Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? *J. Cell. Biochem.* **94**, 890–898 (2005).
4. Small, S., Blair, A. & Levine, M. Regulation of even-skipped stripe 2 in the *Drosophila* embryo. *EMBO J.* **11**, 4047–4057 (1992).
5. Studer, M., Popperl, H., Marshall, H., Kuroiwa, A. & Krumlauf, R. Role of a conserved retinoic acid response element in rhombomere restriction of Hoxb-1. *Science* **265**, 1728–1732 (1994).
6. Arnosti, D. N., Barolo, S., Levine, M. & Small, S. The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. *Development* **122**, 205–214 (1996).
7. Halfon, M. S. *et al.* Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell* **103**, 63–74 (2000).
8. Yuh, C. H., Bolouri, H. & Davidson, E. H. Cis-regulatory logic in the endo16 gene: switching from a specification to a differentiation mode of control. *Development* **128**, 617–629 (2001).
9. Knirr, S. & Frasch, M. Molecular integration of inductive and mesoderm-intrinsic inputs governs even-skipped enhancer activity in a subset of pericardial and dorsal muscle progenitors. *Dev. Biol.* **238**, 13–26 (2001).
10. Oliveri, P., Carrick, D. M. & Davidson, E. H. A regulatory gene network that directs micromere specification in the sea urchin embryo. *Dev. Biol.* **246**, 209–228 (2002).
11. Davidson, B. & Levine, M. Evolutionary origins of the vertebrate heart: Specification of the cardiac lineage in *Ciona intestinalis*. *Proc. Natl Acad. Sci. USA* **100**, 11469–11473 (2003).
12. Hadchouel, J. *et al.* Analysis of a key regulatory region upstream of the Myf5 gene reveals multiple phases of myogenesis, orchestrated at each site by a combination of elements dispersed throughout the locus. *Development* **130**, 3415–3426 (2003).
13. Lee, H. H. & Frasch, M. Nuclear integration of positive Dpp signals, antagonistic Wg inputs and mesodermal competence factors during *Drosophila* visceral mesoderm induction. *Development* **132**, 1429–1442 (2005).
14. Zinzen, R. P., Senger, K., Levine, M. & Papatsenko, D. Computational models for neurogenic gene expression in the *Drosophila* embryo. *Curr. Biol.* **16**, 1358–1365 (2006).
15. Rothbacher, U., Bertrand, V., Lamy, C. & Lemaire, P. A combinatorial code of maternal GATA, Ets and  $\beta$ -catenin-TCF transcription factors specifies and patterns the early ascidian ectoderm. *Development* **134**, 4023–4032 (2007).
16. Sandmann, T. *et al.* A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development. *Dev. Cell* **10**, 797–807 (2006).
17. Zeitlinger, J. *et al.* Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev.* **21**, 385–390 (2007).
18. Sandmann, T. *et al.* A core transcriptional network for early mesoderm development in *Drosophila melanogaster*. *Genes Dev.* **21**, 436–449 (2007).
19. Jakobsen, J. S. *et al.* Temporal ChIP-on-chip reveals Biniou as a universal regulator of the visceral muscle transcriptional network. *Genes Dev.* **21**, 2448–2460 (2007).
20. Li, X. Y. *et al.* Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol.* **6**, e27 (2008).
21. Vokes, S. A., Ji, H., Wong, W. H. & McMahon, A. P. A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. *Genes Dev.* **22**, 2651–2663 (2008).
22. Visel, A. *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858 (2009).
23. Davidson, E. H. *The Regulatory Genome—Gene Regulatory Networks In Development and Evolution* 2nd edn (Elsevier Publishers, 2006).
24. MacArthur, S. *et al.* Developmental roles of 21 *Drosophila* transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions. *Genome Biol.* **10**, R80 (2009).
25. Bintu, L. *et al.* Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.* **15**, 116–124 (2005).
26. Janssens, H. *et al.* Quantitative and predictive model of transcriptional control of the *Drosophila melanogaster* even-skipped gene. *Nature Genet.* **38**, 1159–1165 (2006).
27. Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U. & Gaul, U. Predicting expression patterns from regulatory sequence in *Drosophila* segmentation. *Nature* **451**, 535–540 (2008).
28. Baylies, M. K. & Bate, M. twist: a myogenic switch in *Drosophila*. *Science* **272**, 1481–1484 (1996).
29. Yin, Z., Xu, X. L. & Frasch, M. Regulation of the Twist target gene *tinman* by modular cis-regulatory elements during early mesoderm development. *Development* **124**, 4971–4982 (1997).
30. Azpiazu, N. & Frasch, M. tinman and bagpipe: two homeo box genes that determine cell fates in the dorsal mesoderm of *Drosophila*. *Genes Dev.* **7** (7B), 1325–1340 (1993).
31. Bour, B. A. *et al.* *Drosophila* MEF2, a transcription factor that is essential for myogenesis. *Genes Dev.* **9**, 730–741 (1995).
32. Lilly, B., Galewsky, S., Firulli, A. B., Schulz, R. A. & Olson, E. N. D-MEF2: a MADS box transcription factor expressed in differentiating mesoderm and muscle cell lineages during *Drosophila* embryogenesis. *Proc. Natl Acad. Sci. USA* **91**, 5662–5666 (1994).
33. Zaffran, S., Kuchler, A., Lee, H. H. & Frasch, M. binou (FoxF), a central component in a regulatory network controlling visceral mesoderm development and midgut morphogenesis in *Drosophila*. *Genes Dev.* **15**, 2900–2915 (2001).
34. Furlong, E. E. Integrating transcriptional and signalling networks during muscle development. *Curr. Opin. Genet. Dev.* **14**, 343–350 (2004).
35. Sink, H. *Muscle Development in Drosophila* (Birkhäuser, 2006).
36. Liu, Y. H. *et al.* A systematic analysis of Tinman function reveals Eya and JAK-STAT signaling as essential regulators of muscle development. *Dev. Cell* **16**, 280–291 (2009).
37. Ji, H. & Wong, W. H. TileMap: create chromosomal map of tiling array hybridizations. *Bioinformatics* **21**, 3629–3636 (2005).
38. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-wide mapping of *in vivo* protein-DNA interactions. *Science* **316**, 1497–1502 (2007).
39. Reiss, D. J., Facciotti, M. T. & Baliga, N. S. Model-based deconvolution of genome-wide DNA binding. *Bioinformatics* **24**, 396–403 (2008).
40. Schwartz, Y. B. *et al.* Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. *Nature Genet.* **38**, 700–705 (2006).
41. Cripps, R. M. *et al.* The myogenic regulatory gene *Mef2* is a direct target for transcriptional activation by Twist during *Drosophila* myogenesis. *Genes Dev.* **12**, 422–434 (1998).
42. Cripps, R. M., Zhao, B. & Olson, E. N. Transcription of the myogenic regulatory gene *Mef2* in cardiac, somatic, and visceral muscle cell lineages is regulated by a Tinman-dependent core enhancer. *Dev. Biol.* **215**, 420–430 (1999).
43. Cripps, R. M., Lovato, T. L. & Olson, E. N. Positive autoregulation of the Myocyte enhancer factor-2 myogenic control gene during somatic muscle development in *Drosophila*. *Dev. Biol.* **267**, 536–547 (2004).
44. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).
45. Halfon, M. S., Gallo, S. M. & Bergman, C. M. REDfly 2.0: an integrated database of cis-regulatory modules and transcription factor binding sites in *Drosophila*. *Nucleic Acids Res.* **36** (Database issue), D594–D598 (2008).
46. Bischof, J., Maeda, R. K., Hediger, M., Karch, F. & Basler, K. An optimized transgenesis system for *Drosophila* using germ-line-specific  $\phi$ C31 integrases. *Proc. Natl Acad. Sci. USA* **104**, 3312–3317 (2007).
47. Brown, C. D., Johnson, D. S. & Sidow, A. Functional architecture and evolution of transcriptional elements that drive gene coexpression. *Science* **317**, 1557–1560 (2007).
48. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res.* **35** (Database issue), D88–D92 (2007).
49. Choo, B. G. *et al.* Zebrafish transgenic Enhancer TRAP line database (ZETRAP). *BMC Dev. Biol.* **6**, 5 (2006).
50. Sandmann, T., Jakobsen, J. S. & Furlong, E. E. ChIP-on-chip protocol for genome-wide analysis of transcription factor binding in *Drosophila melanogaster* embryos. *Nature Protocols* **1**, 2839–2855 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We are grateful to M. Leptin for providing an independent assessment of the expression patterns driven by tested CRMs. We thank H. Gustafson for fly work, J. de Graaf for array hybridizations, S. Müller for embryo injections, and R. Bourgon for sharing code on signal peak identification. We thank all members of the Furlong laboratory for discussions and comments on the manuscript. This work was supported by a grant to E.E.M.F. and by a fellowship to R.P.Z. from the Human Frontiers Science Program.

**Author Contributions** M.B. performed ChIP experiments. R.P.Z., E.E.M.F. and C.G. generated CAD. R.P.Z. performed transgenic reporter experiments including *in situ* hybridizations and imaging. C.G. performed ChIP data analysis and motif analysis. J.G. devised the statistical and SVM analyses. E.E.M.F., R.P.Z., C.G. and J.G. formulated the hypotheses, designed experiments and wrote the manuscript.

**Author Information** All ChIP data are available in ArrayExpress under accession numbers E-TABM-648, E-TABM-649, E-TABM-650, E-TABM-651 and E-TABM-652, and the array design under A-AFFY-53. The CRM coordinates and transcription factor occupancy is available at <http://furlonglab.embl.de/>. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to E.E.M.F. ([furlong@embl.de](mailto:furlong@embl.de)).



## METHODS

**ChIP-on-chip assays.** For each transcription factor and time point, two independent chromatin immunoprecipitations (ChIPs) were performed using two different antibodies and compared to two independent mock ChIPs using rabbit pre-immune sera (Supplementary Methods). ChIPs were optimized by assaying for enrichment of a known binding site using real-time PCR as previously described<sup>50</sup>, amplified and hybridized to Affymetrix GeneChip *Drosophila* Tiling array 1.0R (Supplementary Methods).

**Detection of transcription-factor-bound regions and peaks.** All bioinformatics analyses were done using BDGP *Drosophila melanogaster* genome version 4 (UCSC dm2)<sup>51</sup> and the Flybase 4.3 genome annotation release<sup>52</sup>. Mapping of the Affymetrix GeneChip *Drosophila* Tiling 1.0R probes to the genome was obtained from the MAT website (<http://liulab.dfci.harvard.edu/MAT/>). Quantile normalization<sup>53</sup> was applied to the four data sets (two ChIP experiments, two mock controls) for each of the fifteen conditions. Significantly enriched genomic regions at each condition were identified as consecutive probes with significantly positive log-ratios of experiment over control using a Hidden Markov Model (HMM)-based algorithm (TileMap<sup>37</sup>). A cutoff on the probe-wise maximum *a posteriori* probability of each region returned by the HMM was determined manually for each data set (Supplementary Table 1). The top 5% regions below cutoff exhibiting more than 75% overlap with one or more above cutoff region (at any of the 15 conditions) were rescued and included in the final 'high-confidence' transcription factor binding profile. For each enriched region, probe intensity peaks were identified as extrema on a smoothed curve of the log<sub>2</sub>-ratio signal<sup>40</sup>. This 'peak height' was used as a quantitative measure of transcription factor enrichment.

**Enrichment of ChIP peaks near motifs or other peaks.** Enrichment for proximal genomic distances was assessed for (1) distances of transcription factor binding peaks to their closest PWM match (Supplementary Fig. 1) and (2) distances of transcription factor binding peaks in one condition to the closest peak in another condition (for example, distances of Twist binding peaks at 4–6 h to the closest Twist binding peak at 2–4 h, see Supplementary Fig. 2 and Fig. 2b). The background distribution of distances was derived assuming uniform distribution over the merged TileMap regions (that is, significantly enriched regions covered by the tiling array). Enrichment over background was defined as the ratio of the frequency in the data set over background frequency and was robustly estimated using a moving average with a 30-bp window. Equi-tailed 95% confidence intervals of the enrichment were estimated by re-sampling 1,000 times the observed distances, with replacement.

**Computing ChIP CRMs from ChIP peaks.** ChIP peaks across all conditions were clustered using a neighbour joining approach with a maximum distance of 200 bp between adjacent peaks. Peak cluster boundaries were extended by 100 bp past the terminal peak position on each side to account for peak position precision. Resulting genomic boundaries define the ChIP CRMs of the ChIP-CRM atlas (Supplementary Table 5).

**Iterative position weight matrix optimization.** Initial position weight matrices (PWMs) were gathered either from the literature or by *de novo* motif discovery using the software RSAT<sup>54</sup> and GAPWM<sup>55</sup> (Supplementary Methods). Positive genomic regions were defined as 200-bp regions centred on ChIP peaks. PWMs were iteratively updated, taking as binding sites the best hits predicted by the PWM of the previous iteration within relevant positive regions. In computing PWMs, binding sites were weighted by their peak height, giving more importance to peaks with stronger signal. To assess sensitivity and specificity, negative regions were randomly chosen from the repeat-masked genome (excluding exons and bound regions as defined by TileMap), matching the positive regions in number and length. Binding site predictions were generated using Patser<sup>56</sup>. For a given score cutoff, a region was regarded as 'positive' if it contains at least one match above cutoff (ROC curves, Supplementary Fig. 5). Final predictions were performed at a score cutoff corresponding to an estimated 40% false discovery rate.

**Conservation of TFBS within bound regions.** For each transcription factor, ChIP CRMs were split into two groups: a group containing ChIP CRMs bound by the considered transcription factor and a group containing ChIP CRMs not bound by that transcription factor. The latter was used as 'background' regions to control for general sequence biases in ChIP CRMs, such as GC content. In addition, random CRM sets matching the transcription-factor-bound regions in number and length were generated by sampling from the repeat-masked genome (excluding exons and bound regions as defined by TileMap). TFBSs were predicted as described above. Conservation of TFBSs was evaluated in two ways. First, the average of the PhastCons<sup>44</sup> score (dm2/phastCons15way from UCSC, <http://hgdownload.cse.ucsc.edu/goldenPath/dm3/phastCons15way/>) over the bases of the TFBS was computed for either the best scoring TFBS prediction (Fig. 2c) or all TFBS predictions in each CRM (Supplementary Fig. 6a). Second, pair-wise alignments between *D. melanogaster* (dm2) and *D. simulans* (droSim1), *D. yakuba* (droYak1), *D. ananassae* (droAna1), *D. pseudoobscura* (dp3) and *D. virilis*

(droVir1) were obtained from UCSC (<http://hgdownload.cse.ucsc.edu/downloads.html#fruitfly>). Best scoring TFBS predictions (Supplementary Fig. 6b–f) for each ChIP CRM and for each random CRM were used to extract the corresponding sequence from each pair-wise alignment (ungapped alignments only). A TFBS prediction was scored as 'conserved' in a particular species if its aligned sequence triggered a match scoring above cutoff (as defined earlier, see Supplementary Fig. 5), or was otherwise scored as 'not conserved' (unaligned TFBSs were also counted as 'not conserved').

**Compiling the CRM Activity Database (CAD).** The CRM Activity Database (CAD) compiles *in vivo* *Drosophila melanogaster* enhancer expression data from REDfly 2.0<sup>45</sup>, literature surveys and our own experiments. CAD is the result of a semi-automated procedure whereby (1) each enhancer activity is manually reviewed and (2) redundancy within and across data sources is eliminated. We reviewed original literature and enhancer expression pattern images and added missing annotations (using the *Drosophila* gross anatomy ontology, <http://www.obofoundry.org/>), focusing on mesoderm and muscle subset expression. We then manually assessed redundancy of overlapping enhancers; in particular, we minimized enhancer boundaries where possible (overlapping enhancers of varying sizes with indistinguishable activities as assayed) and merged enhancers where necessary (extensively overlapping enhancers having distinct activities). Overlapping enhancers not falling into these two categories were left unmodified. Supplementary Table 4 provides all final CAD entries and references to the original enhancers.

**Defining mesodermal and muscle expression patterns.** While generating CAD, we analysed the expression patterns driven by previously characterized CRMs and annotated them using the *Drosophila* gross anatomy ontology (<http://www.obofoundry.org/>), describing the timing and location of expression. Particular attention was paid to expression in the presumptive and unspecified mesoderm (Meso, stage 5–9), and in 3 of its major derivatives (stage 10 onwards)—somatic musculature (SM), visceral musculature (VM), and heart musculature (CM)<sup>35</sup>. To define mesodermal and muscle expression classes, the anatomical vocabulary of CRMs in CAD was mapped to four high-level master terms: 'Meso', 'SM', 'VM' and 'CM' (Supplementary Table 3). VM-annotated CRMs show expression in circular trunk visceral muscle and/or foregut and hindgut visceral muscle. As the development of longitudinal visceral muscle (IVM) is regulated by transcription factors not included in our data sets, CRMs regulating IVM expression were not annotated as 'VM'.

**The CRM training set.** The training set consisted of 310 computed ChIP CRMs that overlap 250 annotated enhancers in CAD (Supplementary Table 8). We restricted this study to combinatorial expression classes that contained at least 9 positive CRMs: 'Meso' only, 'SM' only, 'VM' only, 'VM & SM' and 'Meso & SM'. These expression classes are mutually exclusive. To be included in the training set for 'Meso', 'VM', 'SM', 'Meso & SM' and 'VM & SM' the CRM had to drive expression in the specific tissue(s), or in subdomains of the tissue(s), at one or more stages of development. CRMs were included that drive expression in the defined expression class, regardless of expression in additional tissues (that is, outside Meso, visceral muscle, somatic muscle and heart musculature), such as ectoderm.

**SVM parameter selection.** We used support vector machines (SVM) with a radial basis function kernel, which requires setting two parameters, the penalization coefficient *C* and the kernel precision  $\gamma$ . Optimal parameter setting was achieved by evaluating the performance for several parameter values. Performance of each SVM was evaluated using the area under the curve (AUC) of receiver operating characteristic (ROC) curves obtained with a leave-one-out cross-validation scheme (Supplementary Information and Supplementary File 1). Sensitivity and specificity functions were smoothed using Gaussian kernel density estimates<sup>57</sup>. Once optimal parameters were identified, a final SVM was trained on the full training set and then applied to the full ChIP-CRM atlas. SVM scores and specificity levels for the 8,008 CRMs of the atlas, and for each expression class, are provided in Supplementary Table 9. For each ChIP CRM, we predicted expression if its SVM score corresponded to a specificity greater than 95% in a given expression class. In cases where CRMs scored with higher than 95% specificity in more than one class, we predicted expression corresponding to the class of higher specificity.

**Transgenic reporter assays.** To assay ChIP CRMs for enhancer activity, the genomic regions were placed in front of a minimal promoter driving a *lacZ* reporter gene in pDuo2n-attB (Supplementary Information). All constructs were targeted to chromosomal position 51C via attB/phiC31 mediated integration<sup>46</sup>. Transgenic lines were balanced, homozygosed and tested by *in situ* hybridization (ISH). *lacZ* ISH was performed either colorimetrically, or fluorescently while co-visualizing appropriate marker genes. No reporter gene activity was detected for an empty vector integrated into that genomic position.

**Scoring CRM activity in transgenic reporter assays.** Colorimetric *lacZ* ISH at all stages of embryogenesis were performed to annotate expression patterns driven

by the CRMs (Fig. 5, Supplementary Figs 9a–13a and Supplementary Methods). Annotated CRM activities are compiled in Supplementary Table 12a, b. Two approaches were taken to confirm the tissue-specific activity of all CRMs.

First, we confirmed all *lacZ* reporter-based annotations using multiplex fluorescent ISH followed by confocal imaging. ISHs were performed using antisense probes against appropriate tissue-specific markers (Supplementary Methods), which were imaged in green (mesoderm and visceral muscle), or blue (somatic muscle), together with a probe directed against the *lacZ* reporter driven by the CRM (red). Overlapping *lacZ* and marker gene expression unequivocally demonstrates tissue-specific CRM activity (Supplementary Figs 9b–13b). The CRMs were annotated as driving expression in a tissue if the expression of the *lacZ* reporter overlapped with that of the marker gene specific for that tissue at any stage of development (Supplementary Table 12a).

Second, we asked an external expert on *Drosophila* mesoderm and muscle development to annotate the CRMs' expression independently (M. Leptin, Institute of Genetics, University of Cologne). M. Leptin was presented with the same embryo pictures shown in Supplementary Figs 9–13, but ordered by increasing CRM ID number and without any information about the respective predictions (Supplementary Information and Supplementary Table 12b).

CRMs were scored as correct SVM predictions if they drove expression within the predicted tissue (or a subregion of that tissue) at any stage of development

and therefore show co-expression with the specific tissue marker. CRMs that drove expression in the predicted tissue, but also in another unpredicted mesoderm tissue, were scored as partial. For the two combinatorial expression classes (Meso & SM and VM & SM), CRMs that only drove expression in one of the two predicted tissues were also scored as partial. CRMs were scored as fail if their expression did not overlap marker gene expression as predicted by the SVM.

51. Celniker, S. E. *et al.* Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence. *Genome Biol.* **3**, RESEARCH0079 (2002).
52. Tweedie, S. *et al.* FlyBase: enhancing *Drosophila* Gene Ontology annotations. *Nucleic Acids Res.* **37** (Database issue), D555–D559 (2009).
53. Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
54. Thomas-Chollier, M. *et al.* RSAT: regulatory sequence analysis tools. *Nucleic Acids Res.* **36** (Web Server issue) W119–W127 (2008).
55. Li, L., Liang, Y. & Bass, R. L. GAPWM: a genetic algorithm method for optimizing a position weight matrix. *Bioinformatics* **23**, 1188–1194 (2007).
56. Hertz, G. Z. & Stormo, G. D. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15**, 563–577 (1999).
57. Lloyd, C. J. Using smoothed receiver operating characteristic curves to summarize and compare diagnostic systems. *J. Am. Stat. Assoc.* **93**, 1356–1364 (1998).



# A neutron star with a carbon atmosphere in the Cassiopeia A supernova remnant

Wynn C. G. Ho<sup>1</sup> & Craig O. Heinke<sup>2</sup>

The surface of hot neutron stars is covered by a thin atmosphere. If there is accretion after neutron-star formation, the atmosphere could be composed of light elements (H or He); if no accretion takes place or if thermonuclear reactions occur after accretion, heavy elements (for example, Fe) are expected. Despite detailed searches, observations have been unable to confirm the atmospheric composition of isolated neutron stars<sup>1</sup>. Here we report an analysis of archival observations of the compact X-ray source in the centre of the Cassiopeia A supernova remnant. We show that a carbon atmosphere neutron star (with low magnetic field) produces a good fit to the spectrum. Our emission model, in contrast with others<sup>2–4</sup>, implies an emission size consistent with theoretical predictions for the radius of neutron stars. This result suggests that there is nuclear burning in the surface layers<sup>5,6</sup> and also identifies the compact source as a very young (~330-year-old) neutron star.

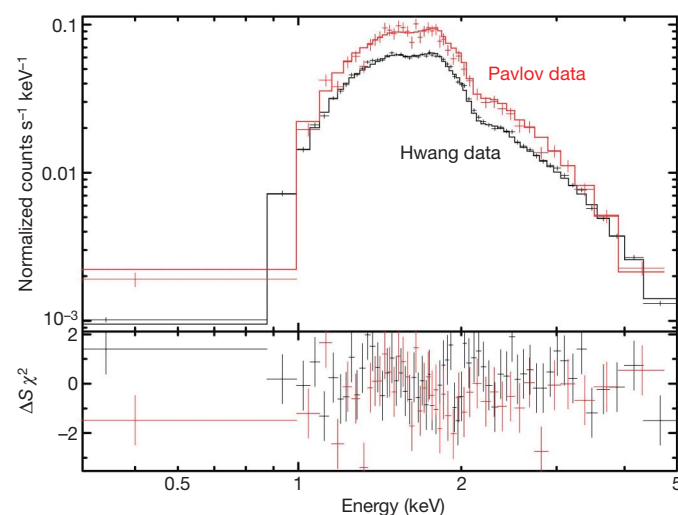
Cassiopeia A is one of the youngest-known supernova remnants in the Milky Way and is at a distance of  $d = 3.4^{+0.3}_{-0.1}$  kiloparsecs (kpc) from the Earth<sup>7</sup>. The supernova that gave rise to the remnant may have been observed by John Flamsteed in 1680 (ref. 8); the implied age coincides with estimates made by studying the expansion of the remnant<sup>9</sup>. Although the supernova remnant is extremely well-studied, the central compact source was only identified in first-light Chandra X-ray observations<sup>10</sup>. (We shall refer to the compact source as Cas A.)

We considered archival Chandra X-ray Observatory data from two studies of Cas A, both using the ACIS-S charge-coupled device which provides spatial and spectral information<sup>11</sup>. A series of Chandra observations, totalling 1 megasecond, was performed in 2004 to study the supernova remnant<sup>12</sup>; these are referred to here as the Hwang data. A shorter (70 kiloseconds) observation in 2006 was designed to study the compact source, here referred to as the Pavlov data<sup>4</sup>. Figure 1 shows the Cas A spectra, as well as our carbon model fit.

We fitted the Hwang and Pavlov data simultaneously with several models: a blackbody, a H atmosphere<sup>13</sup>, and atmospheres composed of pure He, C, N, or O; these are illustrated in Fig. 2. To identify promising models, the mass and radius of the atmosphere models were fixed to the canonical neutron star values of  $M_{\text{NS}} = 1.4M_{\text{Sun}}$  and  $R_{\text{NS}} = 10$  km, where  $M_{\text{Sun}}$  is the solar mass. The normalization factor, which can be interpreted as the fraction of the surface that is emitting X-ray radiation, was left free. The results are given in Table 1.

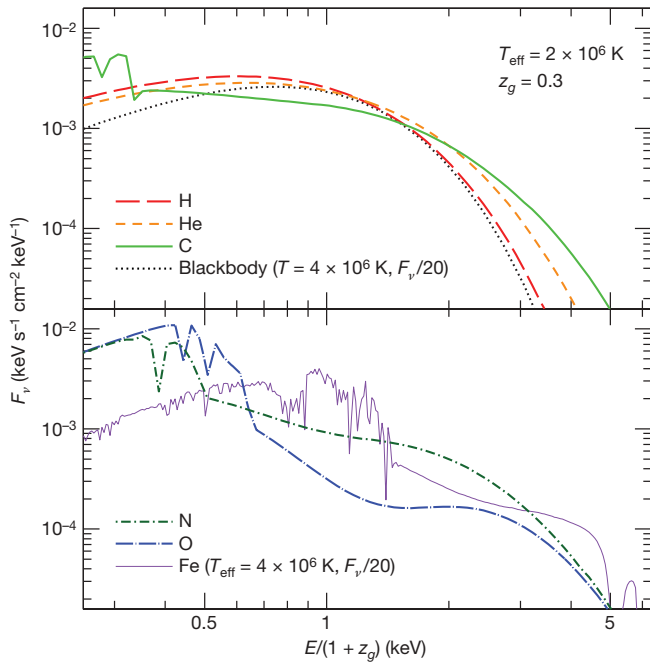
The H, He and C atmospheres provided good fits to the data (somewhat better than the blackbody). However, the derived sizes of the emission region  $R$  for H and He (4 km and 5 km, respectively) are much smaller than the theoretical size of a neutron star  $R_{\text{NS}}$  (~10 km; ref. 14). This would suggest the emission region is a hot spot on the neutron-star surface, which would probably result in

X-ray pulsations as the hot spot rotated with the star. However, these pulsations have not been detected<sup>4,15</sup>. Fits to the data were performed using an additional (temperature) component, for example, a second blackbody or atmosphere spectrum, which produced inferred  $R^{\infty}$  values of ~0.2 km and 2 km for blackbody fits or 0.4 km and 11 km for H atmosphere fits<sup>4</sup>. The high-temperature component has been interpreted as being due to a small hot polar cap, while the low-temperature component is due to emission from the remaining



**Figure 1 | Chandra X-ray spectra of Cas A.** Spectra from the Hwang (black) and Pavlov (red) observations and fits with our C spectral model. Error bars are 1 s.d. The lower panel shows the fit residuals,  $\Delta S\chi^2$  in units of s.d. The Hwang observations place Cas A off-axis (blurring Chandra's point-spread function), and the high count-rate of Cas A distorts the spectrum through detection of multiple photons during one 3.04-s timeframe (known as pile-up<sup>24</sup>). The Pavlov observation is performed in a special instrument configuration to reduce the timeframe to 0.34 s (thus reducing pile-up) and places Cas A at the position of best focus<sup>4</sup>. We used the level 2 event files provided by the Chandra Data Archive and performed data reduction and analysis with CIAO 4.1 and XSPEC 12.4.0. Owing to the slightly distorted shape of the Cas A point-spread function in the Hwang data, we extracted the source spectra using an elliptical region of dimensions 1.23 by 1.72 arcsec, rotated to match the position angle of the point-spread function ellipticity, and the background from a surrounding annulus of 2.19 to 4.37 arcsec. Alternative extraction methods produce similar results. We combine the spectra and responses to make a single, deep spectrum of Cas A. For the Pavlov data, we extract the Cas A spectrum and the nearby background following the procedure of ref. 4. The spectra are binned to achieve at least 1,000 counts per bin for the Hwang data and 140 counts per bin for the Pavlov data. X-ray and bolometric luminosities are  $L_{\text{X}}(0.5\text{--}10\text{ keV}) = 4.3 \times 10^{33} \text{ ergs s}^{-1}$  and  $L_{\text{bol}} = 7.0 \times 10^{33} \text{ ergs s}^{-1}$ , respectively.

<sup>1</sup>School of Mathematics, University of Southampton, Southampton SO17 1BJ, UK. <sup>2</sup>Department of Physics, University of Alberta, Room 238 CEB, 11322-89 Avenue, Edmonton, AB T6G 2G7, Canada.



**Figure 2 | Model atmosphere spectra.** Energy flux  $F$  for atmospheres with H, He, C, N, O, and Fe and a blackbody. The energy has been redshifted by  $1 + z_g = (1 - 2GM_{\text{NS}}/c^2 R_{\text{NS}})^{-1/2} = 1.3$ , where  $z_g$  is the gravitational redshift, and the flux has been scaled by  $(10 \text{ km}/3.4 \text{ kpc})^2$ . Our models are constructed assuming a plane-parallel atmosphere (because the atmosphere thickness of  $\sim 1 \text{ cm}$  is much smaller than the stellar size) that is in hydrostatic and radiative equilibrium with constant gravitational acceleration  $g = (1 + z_g)GM_{\text{NS}}/R_{\text{NS}}^2$ . The efficient separation of light and heavy elements results in atmospheres composed of a single element<sup>23</sup>. The opacities are obtained from tables computed by the Opacity Project<sup>26</sup>. (The energy range of the tables covers  $E/kT \approx 0.07$ – $20$ . When opacities are required beyond this range, we use the  $E^{-3}$ -dependence of free-free and bound-free absorption; this approximation has a minor effect, except for the Fe model, which is shown for illustrative purposes.) Light-element atmospheres generate spectra that are harder than blackbodies (at the same temperature) because of the energy-dependence of the opacity<sup>27,28</sup>, so atmospheric spectral fits result in temperatures that are lower and sizes that are larger compared to those obtained using blackbody spectra. Further details of the atmosphere model construction are given in ref. 29. When  $B \ll 2.35 \times 10^9 \text{ Z}^2 \text{ G}$  ( $\sim 8 \times 10^{10} \text{ G}$  for carbon), magnetic-field effects on the radiation transport and atoms in the atmospheric plasma are negligible<sup>27,28</sup>. Previous works found poor fits to the data using magnetic ( $B \geq 10^{12} \text{ G}$ ) H atmosphere spectra<sup>2,4</sup>, while magnetic mid- $Z$  element spectra are similar to low-magnetic field Fe spectra in that they are blackbody-like in shape and contain many lines<sup>30</sup>.

cooler neutron star surface. However, it is difficult to generate, through anisotropic heat conduction, such small regions of high-temperature contrast on the neutron-star surface<sup>16</sup>. Also, the two temperature regions would again probably produce (undetected)

**Table 1 | X-ray spectral fitting of Cas A**

Atmosphere model	$N_{\text{H}} (10^{22} \text{ cm}^{-2})$	$kT \text{ (eV)}$	Normalization	$\chi^2/\text{d.o.f.}$	Null hypothesis probability (%)
Hydrogen	$1.65^{+0.04}_{-0.05}$	$241^{+7}_{-6}$	$0.18^{+0.03}_{-0.03}$	106.3/99	29
Helium	$1.62^{+0.04}_{-0.05}$	$228^{+9}_{-8}$	$0.22^{+0.05}_{-0.04}$	112.4/99	17
Carbon	$1.73^{+0.04}_{-0.04}$	$155^{+7}_{-6}$	$1.84^{+0.56}_{-0.42}$	105.3/99	31
Nitrogen	1.37	172	1.18	388/99	0
Oxygen	1.03	234	0.20	2,439/99	0
Blackbody model	$1.46^{+0.04}_{-0.04}$	$kT^{\infty} = 387^{+7}_{-6} \text{ eV}$	$R^{\infty} = 1.0^{+0.1}_{-0.1} \text{ km}$	134.2/98	11

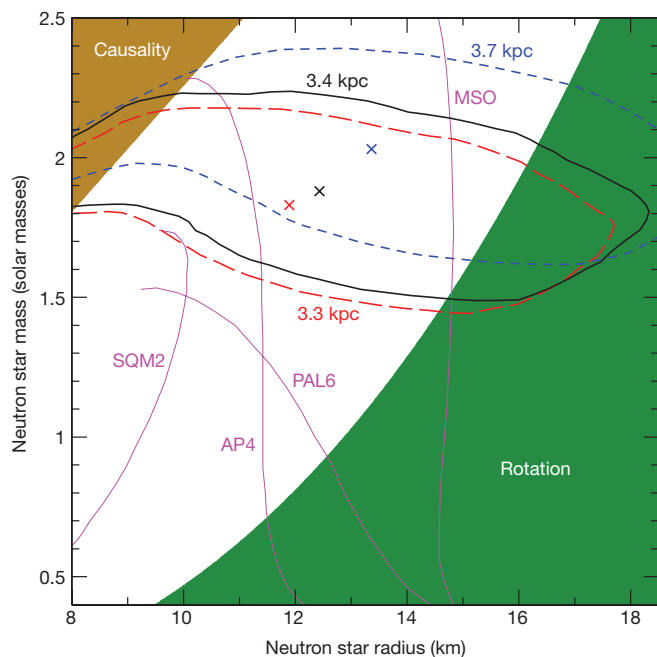
Joint fitting results to the Hwang and Pavlov data with neutron-star atmosphere and blackbody models that are modified by photoelectric absorption  $N_{\text{H}}$  (using the Tuebingen–Boulder absorption routine TBABS with Wilms model abundances in the X-ray spectral fitting package XSPEC; see <http://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/index.html>), dust scattering<sup>23</sup>, and corrections for pile-up<sup>24</sup> (the grade migration parameter of the pile-up algorithm is allowed to float freely and is 0.36 for the best-fit C model). Chandra data with such high statistics reveals systematic uncertainties in the Chandra calibration<sup>15</sup>, so we added a systematic uncertainty of 3% in quadrature. All parameter errors are given at 90% confidence. Errors are not reported when the reduced  $\chi^2 > 2$ . The normalization refers to the fraction of the neutron star surface emitting radiation (for a 10 km stellar radius and 3.4 kpc distance). The null hypothesis probability is the probability that one realization of the model fit to the data would have a reduced  $\chi^2$  greater than that obtained; less than 5% indicates a poor fit.  $T^{\infty} = T/(1 + z_g)$  and  $R^{\infty} = R(1 + z_g)$  are the temperature and radius measured by an observer at infinity and d.o.f. is the number of degrees of freedom.

X-ray pulsations. On the other hand, by considering a C atmosphere, we derived an emission size of  $R \approx 12$ – $15 \text{ km}$  that closely matches the theoretical prediction for the radius of a neutron star<sup>14</sup>,  $R_{\text{NS}} \approx 10$ – $14 \text{ km}$ . In this case, the X-ray observations are detecting emission from the entire stellar surface, and therefore the emission would not necessarily vary as the star rotates (although local temperature differences could cause small brightness fluctuations on top of the bright background). Thus we conclude that Cas A is consistent with a low-magnetic field carbon-atmosphere neutron star of mass  $1.4M_{\text{Sun}}$ , radius 12–15 km, and surface temperature  $T = 1.8 \times 10^6 \text{ K}$ .

Interpreting the size of the emission region to be the true neutron-star radius ( $R = R_{\text{NS}}$ ), we can constrain the neutron-star mass  $M_{\text{NS}}$  and radius  $R_{\text{NS}}$  by using a range of surface gravity models. If we fix  $M_{\text{NS}} = 1.4M_{\text{Sun}}$  and the distance to Cas A as  $d = 3.4 \text{ kpc}$  (ref. 7), then we find a surface effective temperature of  $T_{\text{eff}} = 1.61^{+0.14}_{-0.05} \times 10^6 \text{ K}$  and emission size  $R = 15.6^{+1.3}_{-2.7} \text{ km}$ . Figure 3 shows 90% confidence contours in mass and radius when both are allowed to vary. The contours for distances between 3.3 kpc and 3.7 kpc constrain  $M_{\text{NS}}$  to 1.5– $2.4M_{\text{Sun}}$  and  $R_{\text{NS}}$  to 8–17 km. The mass constraint, significantly above the canonical  $1.4M_{\text{Sun}}$ , suggests a moderately stiff nuclear equation of state<sup>14</sup>.

The emission of neutrinos determines the thermal evolution of a neutron star at ages of less than  $10^5$  years. The particle reactions that contribute to this emission are determined by the state of the matter, which is strongly dependent on the total mass of the star<sup>17,18</sup>. Owing to the small uncertainties in the temperature and age of Cas A, as well as our measured mass, Cas A can be used to constrain matter properties in the stellar interior. More importantly, because the next-youngest neutron stars for which surface thermal emission has been measured have ages exceeding a few thousand years, Cas A (with an age of only about 330 years) serves as a valuable window into the early life of a passively cooling neutron star. Previously, the utility of Cas A in studying the thermal evolution was hindered by the fact that the measured temperature (taken to be from a local hot spot because  $R \ll R_{\text{NS}}$ ) could only be used to set an upper limit on the average surface temperature<sup>2</sup>. However, with our  $T_{\text{eff}}$  and  $L_{\text{bol}}$  (which are surface averages given that  $R = R_{\text{NS}}$ ), detailed comparisons with theoretical models<sup>17–19</sup> can now be made.

The presence of a carbon atmosphere on the surface is probably a consequence of the youth of Cas A. The surface of newly formed neutron stars is uncertain; it is thought to be composed of an element between O and Fe, depending on which layers of the pre-supernova star fall back onto the proto-neutron star<sup>20</sup>; this fallback material could also form light elements through spallation<sup>6</sup>. Over time, the surface accumulates an overlying layer of light elements by accretion from the circumstellar medium (although this can be countered by pulsar-wind excavation<sup>6</sup>). Fits to the thermal radiation from older neutron stars do indeed suggest H or He atmospheres at  $10^4$ – $10^5$  years and heavy-element (blackbody-like) atmospheres at ages exceeding  $10^5$  years<sup>1</sup>. However, the picture is incomplete for young neutron stars



**Figure 3 | Neutron star mass and radius.** 90% confidence contours of  $\chi^2$  around the best-fitting Cas A mass and radius (crosses) for distances of 3.3 kpc (red long-dashed line), 3.4 kpc (thick black solid line), and 3.7 kpc (blue short-dashed line). The upper-left and lower-right regions are excluded by constraints from the requirement of causality and from the fastest-rotating neutron star known<sup>14</sup>. The thin magenta solid lines represent predictions for the masses and radii of neutron stars using different theoretical models (labelled SQM2, AP4, PAL6 and MS0) of the neutron star interior<sup>14</sup>.

(younger than  $10^4$  years). At a depth just below the surface, the temperature (and density) is conducive to nuclear burning ( $T \geq 10^8$  K). Accreted material diffuses to this layer and is rapidly consumed, with depletion of H in under a year and He in under a hundred years<sup>5,6</sup>. This process of diffusive nuclear burning is very temperature-sensitive; as the neutron star cools, the temperature drops below the threshold needed for the material to be burned, and a light-element atmosphere can begin to build up after  $10^4$ – $10^6$  years (ref. 6).

Cas A is the youngest of a class of neutron stars that are located near the centre of supernova remnants, possess steady long-term fluxes and soft thermal spectra, and have no detectable pulsar wind nebulae or radio pulsations<sup>1</sup>. X-ray pulsations have been detected in three members of this class, with periods of tenths of seconds. Timing measurements suggest that they have (dipolar) magnetic fields of  $B \ll 10^{12}$  G (ref. 21). If the magnetic field of Cas A is  $\sim (1\text{--}5) \times 10^{11}$  G, then a spectral feature due to the electron cyclotron resonance may appear in the Chandra energy range, but this has yet to be detected. The lack of a visible pulsar wind nebula and no indication of magnetospheric activity (such as radio or gamma-ray emission or a high-energy hard power-law component), like those seen in classical pulsars (with  $B \geq 10^{12}$  G), also suggest the magnetic field is low. The weak surface magnetic field inferred for this class would have important implications for the neutron-star population. These objects could be representative of the early life of neutron stars before becoming classical pulsars. In this case, either neutron-star magnetic fields develop by a dynamo mechanism, or else a strong field (produced during the collapse of the progenitor star) is buried and has not yet emerged<sup>22</sup>. These objects could also form a distinct group of low-magnetic-field neutron stars, which never manifest pulsar-like behaviour; this would suggest a large population of old, cool, unobserved sources. Finally, we note that the detection of the hydrogen-like C edge at about 0.45 keV (unredshifted), in similar sources with less intervening gas and dust than Cas A has, would not only provide further evidence for carbon atmospheres but also give a measurement of the gravitational redshift.

Received 21 July; accepted 15 September 2009.

- Pavlov, G. G., Zavlin, V. E. & Sanwal, D. In *Proc. WE-Heraeus Seminar: Neutron Stars, Pulsars and Supernova Remnants* (eds Becker, W., Lesch, H., & Trümper, J.) Vol. 270, 273–286 (MPE Report 278, MPI, 2002).
- Pavlov, G. G., Zavlin, V. E., Aschenbach, B., Trümper, J. & Sanwal, D. The compact central object in Cassiopeia A: a neutron star with hot polar caps or a black hole? *Astrophys. J.* **531**, L53–L56 (2000).
- Chakrabarty, D., Pivovarov, M. J., Hernquist, L. E., Heyl, J. S. & Narayan, R. The central X-ray point source in Cassiopeia A. *Astrophys. J.* **548**, 800–810 (2001).
- Pavlov, G. G. & Luna, G. J. M. A dedicated Chandra ACIS observation of the central compact object in the Cassiopeia A supernova remnant. *Astrophys. J.* **703**, 910–921 (2009).
- Rosen, L. C. Hydrogen and helium abundances in neutron-star atmospheres. *Astrophys. Space Sci.* **1**, 372–387 (1968).
- Chang, P. & Bildsten, L. Evolution of young neutron star envelopes. *Astrophys. J.* **605**, 830–839 (2004).
- Reed, J. E., Hester, J. J., Fabian, A. C. & Winkler, P. F. The three-dimensional structure of the Cassiopeia A supernova remnant. I. The spherical shell. *Astrophys. J.* **440**, 706–721 (1995).
- Ashworth, W. B. A probable Flamsteed observation of the Cassiopeia A supernova. *J. Hist. Astron.* **11**, 1–9 (1980).
- Fesen, R. A. et al. The expansion asymmetry and age of the Cassiopeia A supernova remnant. *Astrophys. J.* **645**, 283–292 (2006).
- Tananbaum, H. Cassiopeia A. *IAU Circ.* **7246** (1999).
- Garmire, G. P., Bautz, M. W., Ford, P. G., Nousek, J. A. & Ricker, G. R. Advanced CCD imaging spectrometer (ACIS) instrument on the Chandra X-ray Observatory. *Proc. SPIE* **4851**, 28–44 (2003).
- Hwang, U. et al. A million second Chandra view of Cassiopeia A. *Astrophys. J.* **615**, L117–L120 (2004).
- Heinke, C. O., Rybicki, G. B., Narayan, R. & Grindlay, J. E. A hydrogen atmosphere spectral model applied to the neutron star X7 in the globular cluster 47 Tucanae. *Astrophys. J.* **644**, 1090–1103 (2006).
- Lattimer, J. M. & Prakash, M. Neutron star observations: prognosis for equation of state constraints. *Phys. Rep.* **442**, 109–165 (2007).
- Murray, S. S., Ransom, S. M., Juda, M., Hwang, U. & Holt, S. S. Is the compact source at the center of Cassiopeia A pulsed? *Astrophys. J.* **566**, 1039–1044 (2002).
- Greenstein, G. & Hartke, G. J. Pulselike character of blackbody radiation from neutron stars. *Astrophys. J.* **271**, 283–293 (1983).
- Yakovlev, D. G. & Pethick, C. J. Neutron star cooling. *Annu. Rev. Astron. Astrophys.* **42**, 169–210 (2004).
- Page, D., Geppert, U. & Weber, F. The cooling of compact stars. *Nucl. Phys. A* **777**, 497–530 (2006).
- Tsuruta, S. et al. Thermal evolution of hyperon-mixed neutron stars. *Astrophys. J.* **691**, 621–632 (2009).
- Woosley, S. E., Heger, A. & Weaver, T. A. The evolution and explosion of massive stars. *Rev. Mod. Phys.* **74**, 1015–1071 (2002).
- Gotthelf, E. V. & Halpern, J. P. Discovery of a 112 ms X-ray pulsar in Puppis A: further evidence of neutron stars weakly magnetized at birth. *Astrophys. J.* **695**, L35–L39 (2009).
- Muslimov, A. & Page, D. Delayed switch-on of pulsars. *Astrophys. J.* **440**, L77–L80 (1995).
- Predehl, P., Costantini, E., Hasinger, G. & Tanaka, Y. XMM-Newton observation of the galactic centre—evidence against the X-ray reflection nebulae model? *Astron. Nachr.* **324**, 73–76 (2003).
- Davis, J. E. Event pileup in charge-coupled devices. *Astrophys. J.* **562**, 575–582 (2001).
- Alcock, C. & Illarionov, A. Surface chemistry of stars. I. Diffusion of heavy ions in white dwarf envelopes. *Astrophys. J.* **235**, 534–540 (1980).
- The Opacity Project Team. *The Opacity Project* (<http://cdsweb.u-strasbg.fr/topbase/op.html>) (2009).
- Rajagopal, M. & Romani, R. W. Model atmospheres for low-field neutron stars. *Astrophys. J.* **461**, 327–333 (1996).
- Zavlin, V. E., Pavlov, G. G. & Shibano, Yu. A. Model neutron star atmospheres with low magnetic fields. *Astron. Astrophys.* **315**, 141–152 (1996).
- Ho, W. C. G. & Lai, D. Atmospheres and spectra of strongly magnetized neutron stars. *Mon. Not. R. Astron. Soc.* **327**, 1081–1096 (2001).
- Mori, K. & Ho, W. C. G. Modelling mid-Z element atmospheres for strongly magnetized neutron stars. *Mon. Not. R. Astron. Soc.* **377**, 905–919 (2007).

**Acknowledgements** W.C.G.H. thanks N. Badnell, P. Chang, and D. Lai for discussions. W.C.G.H. appreciates the use of the computer facilities at the Kavli Institute for Particle Astrophysics and Cosmology. W.C.G.H. acknowledges support from the Science and Technology Facilities Council (STFC) in the United Kingdom. C.O.H. acknowledges support from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

**Author Contributions** W.C.G.H. calculated the new models and wrote the manuscript. C.O.H. reduced the data, fitted the models to the data, and contributed to the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to W.C.G.H. ([wynnho@slac.stanford.edu](mailto:wynnho@slac.stanford.edu)) or C.O.H. ([cheinke@phys.ualberta.ca](mailto:cheinke@phys.ualberta.ca)).



## LETTERS

# A quantum gas microscope for detecting single atoms in a Hubbard-regime optical lattice

Waseem S. Bakr<sup>1</sup>, Jonathon I. Gillen<sup>1</sup>, Amy Peng<sup>1</sup>, Simon Fölling<sup>1</sup> & Markus Greiner<sup>1</sup>

Recent years have seen tremendous progress in creating complex atomic many-body quantum systems. One approach is to use macroscopic, effectively thermodynamic ensembles of ultracold atoms to create quantum gases and strongly correlated states of matter, and to analyse the bulk properties of the ensemble. For example, bosonic and fermionic atoms in a Hubbard-regime optical lattice<sup>1–5</sup> can be used for quantum simulations of solid-state models<sup>6</sup>. The opposite approach is to build up microscopic quantum systems atom-by-atom, with complete control over all degrees of freedom<sup>7–9</sup>. The atoms or ions act as qubits and allow the realization of quantum gates, with the goal of creating highly controllable quantum information systems. Until now, the macroscopic and microscopic strategies have been fairly disconnected. Here we present a quantum gas ‘microscope’ that bridges the two approaches, realizing a system in which atoms of a macroscopic ensemble are detected individually and a complete set of degrees of freedom for each of them is determined through preparation and measurement. By implementing a high-resolution optical imaging system, single atoms are detected with near-unity fidelity on individual sites of a Hubbard-regime optical lattice. The lattice itself is generated by projecting a holographic mask through the imaging system. It has an arbitrary geometry, chosen to support both strong tunnel coupling between lattice sites and strong on-site confinement. Our approach can be used to directly detect strongly correlated states of matter; in the context of condensed matter simulation, this corresponds to the detection of individual electrons in the simulated crystal. Also, the quantum gas microscope may enable addressing and read-out of large-scale quantum information systems based on ultracold atoms.

In Hubbard-regime optical lattice systems, an atomic quantum gas resides in the lowest Bloch band of a multi-dimensional array of lattice sites<sup>10</sup>. Strongly correlated quantum states, such as bosonic and fermionic Mott insulator states, have been created in such lattices<sup>2,4,5</sup>, and the realization of quantum magnetism and *d*-wave superfluidity is being actively pursued<sup>6,11,12</sup>. Experiments of this type have for the most part relied on measuring ensemble properties, such as global coherence and compressibility. Creating the possibility of probing the quantum gas with single atom/single lattice site resolution, in contrast, would allow experimenters to access the quantum gas on a single ‘qubit’ level and measure the particle–particle correlation functions that characterize strongly correlated quantum states.

The quantum gas microscope provides exactly this capability through an unprecedented combination of resolution and sensitivity. It is based on a number of innovations, such as solid immersion microscopy for cold atoms, a two-dimensional (2D) optical lattice in close vicinity to an optical surface, and the use of incoherent lattice light. The microscope enables the detection of single atoms with near-unity fidelity on single lattice sites of a short period optical lattice in the Hubbard regime. Previously, site-resolved optical imaging of single atoms has been demonstrated in lattices with large

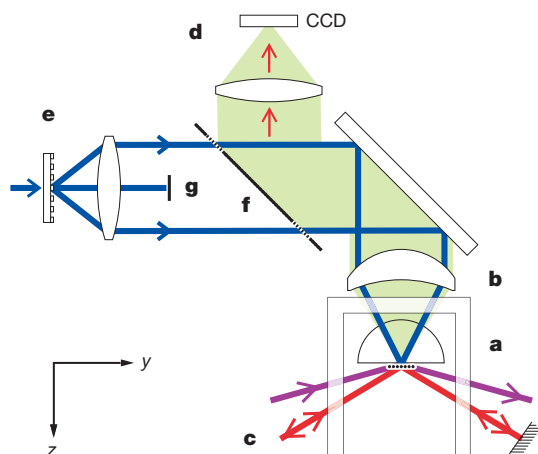
spacings (5  $\mu\text{m}$  period)<sup>13</sup> and in sparsely populated one-dimensional arrays<sup>14</sup>. Imaging of 2D arrays of ‘tubes’ with large occupations has been shown for smaller spacings with an electron microscope<sup>15</sup> and optical imaging<sup>16</sup> systems. For the described applications, however, a combination of high fidelity single atom detection and short lattice periods is important, which has not been previously achieved. The central feature of the Hubbard regime is that the tunnel coupling between sites and the on-site interaction must be comparable in strength, and should be larger than the temperature of the system. Small lattice spacings of the order of 500 nm are required to fulfil these requirements while also obtaining dynamical timescales much shorter than the lifetime of the Hubbard system. We demonstrate high fidelity site-resolved single atom detection in such a lattice.

The quantum gas microscope is based on a high aperture optical system, which simultaneously serves to generate lattice potentials and detect the atoms. By placing a 2D quantum gas only a few micrometres away from the front surface of this microscope, we are able to achieve a very high numerical aperture of  $\text{NA} = 0.8$ . As a result, we measure an optical resolution of  $\sim 600$  nm (full-width at half-maximum, FWHM). Unlike typical optical lattice set-ups, we create the lattice potential by directly projecting a spatial light pattern<sup>16–20</sup> onto the atom plane. A lithographically produced holographic mask and our high resolution optics allow us to generate arbitrary potential landscapes with sub-micrometre structures, opening the possibility of realizing a wide range of model Hamiltonians. Here we project a simple cubic lattice with a lattice constant of 640 nm, comparable to typical Hubbard-type lattice experiments. After loading the lattice, we can directly read out all (up to tens of thousands) lattice sites individually by imaging the light scattered by the atoms. The 2D geometry of the system enables imaging without line-of-sight integration, allowing for direct, reconstruction-free detection of densities<sup>21</sup> in our system on a single atom level.

The central part of the set-up is the high resolution optical imaging system integrated with a 2D atom trap (Fig. 1). The imaging system consists of a long working distance microscope objective located outside the vacuum chamber which covers a numerical aperture of  $\text{NA} = 0.55$ . As an additional front lens of this imaging system, a hemispherical lens is placed inside the vacuum. With the quantum gas placed only a few micrometres from the superpolished flat bottom surface of the hemisphere, a ‘solid immersion’<sup>22</sup> effect occurs, which increases the numerical aperture by the index of refraction of the hemisphere lens to  $\text{NA} = 0.8$ , yielding a diffraction limit of 500 nm at an imaging wavelength of 780 nm. The quantum gas of <sup>87</sup>Rb atoms is created in a hybrid surface trap based on evanescent and standing-wave light fields<sup>23</sup>. Positioned between 1.5 and 3  $\mu\text{m}$  from the surface, the quantum gas is deep in the 2D regime, with trap frequencies of 6 kHz in the vertical direction and 20 Hz in the horizontal plane.

The periodic potentials in the 2D plane are created by using the microscope optics to make a direct projection of a lithographically produced periodic mask that contains the lattice structure in the form

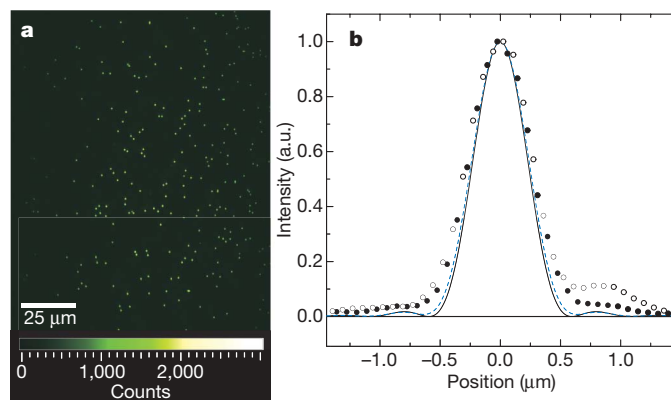
<sup>1</sup>Harvard-MIT Center for Ultracold Atoms and Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA.



**Figure 1 | Diagram of the quantum gas microscope.** The two-dimensional atom sample (a) is located a few micrometres below the lower surface of a hemispherical lens inside the vacuum chamber. This lens serves to increase the numerical aperture (NA) of the objective lens outside the vacuum (b) by the index of refraction, from  $NA = 0.55$  to  $NA = 0.8$ . The atoms are illuminated from the side by the molasses beams (c) and the scattered fluorescence light is collected by the objective lens and projected onto a CCD camera (d). A 2D optical lattice is generated by projecting a periodic mask (f) onto the atoms through the same objective lens via a beam splitter (g). The mask is a periodic phase hologram, and a beam stop (e) blocks the residual zeroth order, leaving only the first orders to form a sinusoidal potential.

of a phase hologram. This is in contrast to conventional optical lattice experiments in which lattice potentials are created by superimposing separate laser beams to create optical standing waves. The advantage of the new method is that the geometry of the lattice is directly given by the pattern on the mask. The imaged light pattern, and hence the potential landscape, can be arbitrary within the limits set by the available imaging aperture and by polarization effects that can arise due to the large aperture imaging beyond the paraxial limit. Here, we create blue detuned square lattice potentials with a periodicity  $a = 640$  nm and an overall Gaussian envelope. A major additional advantage is the fact that the lattice geometry is not dependent on the wavelength<sup>20</sup>, apart from diffraction limits and chromatic aberrations in the lens for large wavelength changes. This allows us to use spectrally wide ‘white’ light with a short coherence length to reduce unwanted disorder from stray light interference. With a light source centred around 758 nm, we generate a conservative lattice potential with a lattice depth of up to  $35 E_{\text{rec}}$ , where  $E_{\text{rec}} = \hbar^2/8ma^2$  is the recoil energy of the effective lattice wavelength, with  $m$  the mass of  $^{87}\text{Rb}$ .

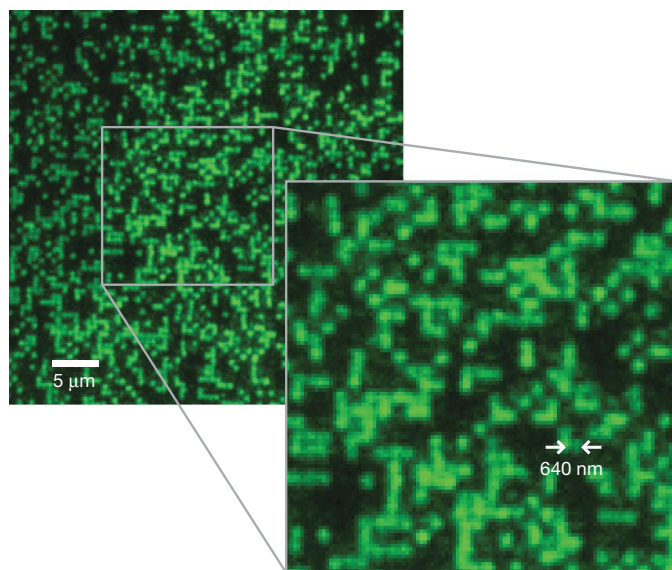
The projection method also enables us to dynamically change the wavelength of the lattice light without changing the lattice geometry. This is important, as we strongly increase the lattice depth for site-resolved imaging in order to suppress diffusion of the atoms between sites due to recoil heating by the imaging light<sup>13</sup>. For this, we switch the light in the 2D lattice and the vertical standing wave to near-resonant narrow band light, increasing the lattice depth to  $5,500 E_{\text{rec}}$  (to 380  $\mu\text{K}$ ). The main use of the microscope set-up is the collection of fluorescence light and high-resolution imaging of the atoms. With the atoms pinned to the deep lattice, we illuminate the sample with red detuned near-resonant light in an optical molasses configuration, which simultaneously provides sub-Doppler cooling<sup>24,25</sup>. Figure 2 shows a typical image obtained by loading the lattice with a very dilute cloud, showing the response of individual atoms. The spot function of a single atom can be directly obtained from such images. We measure a typical single atom emission FWHM size as 570 nm and 630 nm along the  $x$  and  $y$  direction, respectively, which is close to the theoretical minimum value of  $\sim 520$  nm (Fig. 3). This minimum is given by the diffraction limit from the objective combined with the finite size of the camera



**Figure 2 | Imaging single atoms.** a, Field of view with sparse site occupation. b, Response of a single atom, derived from sparse images: shown are horizontal (filled circles) and vertical (open circles) profiles through the centre of the image generated by a single atom. The black line shows the expected Airy function for a perfect imaging system with a numerical aperture of 0.8. The blue dashed line denotes the profile expected from a single atom, taking into account only the finite width of the CCD pixels and the finite extension of the probability distribution of the atom's location. The data are from the responses of 20 atoms in different locations within the field of view which have been precisely superimposed by subpixel shifting before averaging.

pixels and the expected extent of the atom's on-site probability distribution within the lattice site during the imaging process. As the same high-resolution optics are used to generate both the lattice and the image of the atoms on the CCD camera, the imaging system is very stable with respect to the lattice, which is important for single-site addressing<sup>26</sup>. The observed drifts in the 2D plane are very low, less than 10% of the lattice spacing in one hour with shot to shot fluctuations of less than 15% r.m.s.

Pair densities within multiply occupied lattice sites are very high due to the strong confinement in the lattice. When resonantly illuminated, such pairs undergo light assisted collisions and leave the trap within a time of the order of 100  $\mu\text{s}$ , long before they emit sufficient photons to be detected<sup>27</sup>. Therefore the remaining number

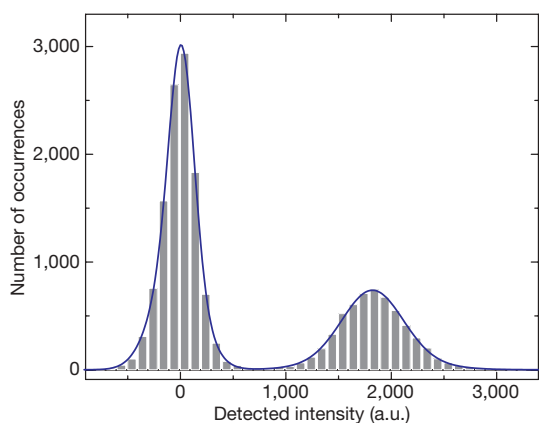


**Figure 3 | Site-resolved imaging of single atoms on a 640-nm-period optical lattice, loaded with a high density Bose-Einstein condensate.** Inset, magnified view of the central section of the picture. The lattice structure and the discrete atoms are clearly visible. Owing to light-assisted collisions and molecule formation on multiply occupied sites during imaging, only empty and singly occupied sites can be seen in the image.

of atoms per site is equal to the parity of the original atom number before illumination, as long as the initial occupation is small. For our molasses parameters, the collected number of photons can be up to  $2 \times 10^4$  per atom per second, and the exposure times are typically between 200 and 1,000 ms, limited by the loss of single atoms from the trap which reduces the detection fidelity. The  $1/e$  lifetime is  $\sim 30$  s, which is consistent with loss due to collisions with hot atoms in the background gas.

Figure 3 shows an image obtained by loading a dense Bose–Einstein condensate. The fast ramp-up of the pinning lattice within 1.5 ms switches off tunnelling and projects the superfluid state wavefunction onto Poisson distributed on-site occupations with more than one atom per lattice site in the centre of the trap. Owing to the removal of pairs the occupation detected is lowered, typically to 42%. The images are analysed by identifying the lattice geometry and fitting point spread functions (obtained separately by analysing images from sparsely filled lattices) to each lattice point. As the background signal is weak and smooth due to the 2D geometry, we thus obtain the total number of scattered photons per lattice site as a simple way of determining the presence of an atom. Figure 4 shows the histogram of photon counts for the central region of several images with an average filling of 34%. For these pictures with long exposure times, the fidelity of identifying atoms at a given lattice site is 98%, limited by the losses occurring during the integration time. To verify that the atom distribution is preserved during imaging, we have recorded sequences of consecutive images spanning a total detection period of several seconds, during which no significant hopping occurs.

In conclusion, we have demonstrated site resolved high fidelity imaging of thousands of individual atoms in a Hubbard-regime optical lattice. A complete set of degrees of freedom of each of them is determined through preparation and measurement. Such detection opens many new possibilities. Strongly correlated quantum states, such as the Mott insulator and antiferromagnetic states, could be directly detected, making it possible to determine results of quantum simulations on a single ‘qubit’ level. In our lattice configuration, the Mott transition is expected at a lattice depth of  $\sim 12 E_{\text{rec}}$  with a tunnel coupling strength  $J$  of 20 Hz at the transition point, similar to the parameters in ref. 21. Single atom detection would then enable the precise measurement of entropy or temperature via the direct detection of defects. Similarly, such imaging is an enabling method for quantum information applications. Based on a Mott insulator state and time-varying state-dependent potentials, maximally entangled cluster states of tens of atoms have been created<sup>28</sup>, which could be used as a resource for quantum computations<sup>29</sup>. For detecting or exploiting this entanglement, however, single-site readout and manipulation are essential. Generating our lattice potentials by direct projection of a 2D



**Figure 4 | Histogram showing the brightness distribution of lattice sites for an exposure time of 1 s.** The left peak corresponds to empty sites (background subtracted), the right peak to those occupied by a single atom. The blue line denotes a fit to the data using a double Gaussian function for each of the two peaks. a.u., arbitrary units.

mask gives us the additional advantage of full control over the lattice geometry on a single-site level. The method can be generalized to create arbitrary potential landscapes, including spin-dependent ones, for realizing a large class of model Hamiltonians. Local manipulation could also enable new cooling schemes<sup>30</sup> for reaching temperatures low enough to study quantum magnetism and  $d$ -wave superfluidity.

## METHODS SUMMARY

A 2D Bose–Einstein condensate of  $^{87}\text{Rb}$  is created by loading a three-dimensional (3D) cloud into a hybrid surface trap, described in detail in ref. 23. The 2D trap is formed by using a magnetic gradient to push a spherical condensate against a repulsive dipole potential due to an evanescent wave from a blue-detuned beam totally internally reflected inside the hemispherical lens. The resulting pancake-shaped condensate is then loaded into a single well of a one-dimensional (1D) far off-resonant lattice with nodal planes parallel to the glass spaced at  $1.5 \mu\text{m}$ . The final geometry has axial and transverse confinements of 6 kHz and 20 Hz, respectively.

A square lattice light pattern created using holographic masks is projected onto the atom plane through the high resolution objective. The lattice spacing in the atom plane is 640 nm. By an appropriate choice of polarizations and frequencies, the lattice light is linearly polarized everywhere. This is essential to avoid effective magnetic fields in the case of near resonant light that can interfere with polarization gradient cooling in the imaging molasses. We create a conservative lattice by illuminating the holograms with blue detuned broadband light at 22 nm from the D2 transition. For imaging, we switch on a superposed pinning lattice with a depth of 380  $\mu\text{K}$ , generated using the same holograms with a separate narrow band light source detuned 32 GHz from the D1 line. The atoms are then illuminated in a molasses configuration at 80 MHz from the D2 line and imaged by collecting the scattered photons onto a CCD. The photon collection efficiency is  $\sim 10\%$ .

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 20 July; accepted 3 September 2009.

- Jaksch, D., Bruder, C., Cirac, J. I., Gardiner, C. W. & Zoller, P. Cold bosonic atoms in optical lattices. *Phys. Rev. Lett.* **81**, 3108–3111 (1998).
- Greiner, M., Mandel, O., Esslinger, T., Hänsch, T. & Bloch, I. Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms. *Nature* **415**, 39–44 (2002).
- Köhl, M., Moritz, H., Stöferle, T., Günter, K. & Esslinger, T. Fermionic atoms in a three dimensional optical lattice: observing Fermi surfaces, dynamics, and interactions. *Phys. Rev. Lett.* **94**, 080403 (2005).
- Jördens, R., Strohmaier, N., Günter, K., Moritz, H. & Esslinger, T. A Mott insulator of fermionic atoms in an optical lattice. *Nature* **455**, 204–207 (2008).
- Schneider, U. *et al.* Metallic and insulating phases of repulsively interacting fermions in a 3D optical lattice. *Science* **322**, 1520–1525 (2008).
- Bloch, I., Dalibard, J. & Zwierger, W. Many-body physics with ultracold gases. *Rev. Mod. Phys.* **80**, 885–964 (2008).
- Schrader, D. *et al.* Neutral atom quantum register. *Phys. Rev. Lett.* **93**, 150501 (2004).
- Blatt, R. & Wineland, D. J. Entangled states of trapped atomic ions. *Nature* **453**, 1008–1014 (2008).
- Urban, E. *et al.* Observation of Rydberg blockade between two atoms. *Nature Phys.* **5**, 110–114 (2009).
- Jaksch, D. & Zwierger, W. The cold atom Hubbard toolbox. *Ann. Phys.* **315**, 52–79 (2005).
- Duan, L.-M., Demler, E. & Lukin, M. D. Controlling spin exchange interactions of ultracold atoms in optical lattices. *Phys. Rev. Lett.* **91**, 090402 (2003).
- Trotzky, S. *et al.* Time-resolved observation and control of superexchange interactions with ultracold atoms in optical lattices. *Science* **319**, 295–299 (2008).
- Nelson, K., Li, X. & Weiss, D. Imaging single atoms in a three-dimensional array. *Nature Phys.* **3**, 556–560 (2007).
- Karski, M. *et al.* Nearest-neighbor detection of atoms in a 1D optical lattice by fluorescence imaging. *Phys. Rev. Lett.* **102**, 053001 (2009).
- Gericke, T., Würtz, P., Reitz, D., Langen, T. & Ott, H. High-resolution scanning electron microscopy of an ultracold quantum gas. *Nature Phys.* **4**, 949–953 (2008).
- Itah, A. *et al.* Direct observation of number squeezing in an optical lattice. Preprint at (<http://arXiv.org/abs/0903.3282>) (2009).
- Grier, D. G. A revolution in optical manipulation. *Nature* **424**, 810–816 (2003).
- Newell, R., Sebby, J. & Walker, T. G. Dense atom clouds in a holographic atom trap. *Opt. Lett.* **28**, 1266–1268 (2003).
- Bergamini, S. *et al.* Holographic generation of microtrap arrays for single atoms by use of a programmable phase modulator. *J. Opt. Soc. Am. B* **21**, 1889–1894 (2004).



20. Brickman Soderberg, K., Gemelke, N. & Chin, C. Ultracold molecules: vehicles to scalable quantum information processing. *N. J. Phys.* **11**, 055022 (2009).
21. Gemelke, N., Zhang, X., Hung, C. & Chin, C. *In situ* observation of incompressible Mott-insulating domains in ultracold atomic gases. *Nature* **460**, 995–998 (2009).
22. Mansfield, S. M. & Kino, G. S. Solid immersion microscope. *Appl. Phys. Lett.* **57**, 2615–2616 (1990).
23. Gillen, J. I., Bakr, W. S., Peng, A., Unterwadtzer, P., Fölling, S. & Greiner, M. Two-dimensional quantum gas in a hybrid surface trap. *Phys. Rev. A* **80**, 021602(R) (2009).
24. Wineland, D. J., Dalibard, J. & Cohen-Tannoudji, C. Sisyphus cooling of a bound atom. *J. Opt. Soc. Am. B* **9**, 32–42 (1992).
25. Winoto, S. L., DePue, M., Bramall, N. E. & Weiss, D. S. Laser cooling at high density in deep far-detuned optical lattices. *Phys. Rev. A* **59**, R19–R22 (1999).
26. Würtz, P., Langen, T., Gericke, T., Koglbauer, A. & Ott, H. Experimental demonstration of single-site addressability in a two-dimensional optical lattice. *Phys. Rev. Lett.* **103**, 080404 (2009).
27. DePue, M. T., McCormick, C., Winoto, S. L., Oliver, S. & Weiss, D. S. Unity occupation of sites in a 3D optical lattice. *Phys. Rev. Lett.* **82**, 2262–2265 (1999).
28. Mandel, O. *et al.* Controlled collisions for multi-particle entanglement of optically trapped atoms. *Nature* **425**, 937–940 (2003).
29. Raussendorf, R. & Briegel, H. J. A one-way quantum computer. *Phys. Rev. Lett.* **86**, 5188–5191 (2001).
30. Bernier, J. *et al.* Cooling fermionic atoms in optical lattices by shaping the confinement. *Phys. Rev. A* **79**, 061601 (2009).

**Acknowledgements** We are grateful for discussions with D. Weiss and V. Vuletic and thank P. Unterwadtzer, E. Su and J. Brachmann for experimental assistance during the early stage of the experiment. We thank D. Weiss for sharing the lens design of the objective lens. This work was funded by grants from the NSF, AFOSR MURI, DARPA, an Alfred P. Sloan Fellowship to M.G., and an NSF Fellowship to J.I.G.

**Author Contributions** All authors contributed to the design and building of the set-up and taking of the data. W.S.B. and S.F. analysed the data and W.S.B., S.F. and M.G. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to M.G. ([greiner@physics.harvard.edu](mailto:greiner@physics.harvard.edu)).

## METHODS

**Hybrid surface trap.** The hybrid surface trap, described in detail elsewhere<sup>23</sup>, is loaded in a two-step sequence. First, a Bose–Einstein condensate created in a spherical magnetic trap is moved against the flat, superpolished glass surface. Close to the glass, it experiences a repulsive dipole potential due to an evanescent wave from a 767 nm blue detuned beam with a spectral width of 2 nm which is totally internally reflected inside the glass. This results in tight axial confinement with trap frequencies of typically 900 Hz. The weak axial confinement of 20 Hz is provided by the magnetic trap. In order to increase the axial confinement to the final value, the 2D gas is transferred into a vertical standing wave generated by a blue-detuned far off-resonant beam reflected at 15° from the trapping surface. Using radio-frequency forced evaporative cooling, we achieve nearly pure condensates in the 2D trap, as verified by time-of-flight imaging and matter wave interference<sup>23</sup>.

**Optical lattice projection.** We create arbitrary potential landscapes by imaging holographic masks onto the atom plane. In this experiment, we use two binary phase holograms to create sinusoidal potentials along the  $x$  and  $y$  axes respectively. The holograms are illuminated with linearly polarized light, polarized perpendicular to the plane of diffraction. The two light paths are combined with a polarizing beam splitter cube. The far off-resonant lattice is created with light from a femtosecond laser, with a spectral width of 3 nm centred at 758 nm. The short coherence length improves the quality of the transverse lattice potential owing to the elimination of uncontrolled interference with stray light<sup>23</sup>. The lattice potential produced by the holograms is given by  $V(x,y) = V_0(\sin^2(kx) + \sin^2(ky))$ , where the periodicity of the lattice is given by  $a = \pi/k = 640$  nm and  $V_0$ , the depth of the potential, can reach up to 35  $E_{\text{rec}}$ . For imaging, we increase the lattice depth to 5,500

$E_{\text{rec}}$  (0.38 mK) by illuminating the lattice holograms with light from a continuous wave Ti:sapphire laser detuned 32 GHz to the blue of the D1 transition of  $^{87}\text{Rb}$ , without changing the lattice geometry. This pinning lattice is linearly polarized everywhere to avoid effective magnetic fields that interfere with the polarization gradient cooling during imaging. This is achieved using the proper choice of polarizations and by introducing frequency differences of at least 80 MHz between lattice axes to time average the interference between them. Near resonant light from the same source is used to simultaneously increase the lattice depth in the axial direction to 3 mK.

**Fluorescence imaging.** After the atoms are pinned, we image the atoms by illuminating them with light detuned 80 MHz to the red of the  $F = 2$  to  $F' = 3$  transition of the D2 line, where  $F$  and  $F'$  denote the hyperfine manifold in the ground and excited state, respectively. The light is in an optical molasses configuration which simultaneously provides illumination and cooling. One beam enters from the  $y$  axis and is reflected off the trapping surface at an angle of 15° and then retroreflected with perpendicular polarization along the same path. This results in polarization and intensity gradients along the  $y$  direction and the vertical. An additional beam enters along the  $x$  axis which generates polarization gradient components along this axis by interference with the retroreflected beam. In addition, to avoid cooling inefficiencies due to low polarization gradients on some lattice sites, we frequency offset the molasses beams by 7 kHz for temporal averaging of the cooling pattern.

The photons scattered by the optical molasses are collected for fluorescence detection of the atoms. The solid angle of the imaging system leads to a collection efficiency of 20%, such that we expect a total photon collection efficiency of ~10% including the quantum efficiency of the CCD camera (Andor Ixon DU888). The effective pixel size in the object plane is 167 nm.

## LETTERS

## Optomechanical crystals

Matt Eichenfield<sup>1</sup>, Jasper Chan<sup>1</sup>, Ryan M. Camacho<sup>1</sup>, Kerry J. Vahala<sup>1</sup> & Oskar Painter<sup>1</sup>

Periodicity in materials yields interesting and useful phenomena. Applied to the propagation of light, periodicity gives rise to photonic crystals<sup>1</sup>, which can be precisely engineered for such applications as guiding and dispersing optical beams<sup>2,3</sup>, tightly confining and trapping light resonantly<sup>4</sup>, and enhancing nonlinear optical interactions<sup>5</sup>. Photonic crystals can also be formed into planar lightwave circuits for the integration of optical and electrical microsystems<sup>6</sup>. In a photonic crystal, the periodicity of the host medium is used to manipulate the properties of light, whereas a phononic crystal uses periodicity to manipulate mechanical vibrations<sup>7–13</sup>. As has been demonstrated in studies of Raman-like scattering in epitaxially grown vertical cavity structures<sup>14</sup> and photonic crystal fibres<sup>15</sup>, the simultaneous confinement of mechanical and optical modes in periodic structures can lead to greatly enhanced light–matter interactions. A logical next step is thus to create planar circuits that act as both photonic and phononic crystals<sup>16</sup>: optomechanical crystals. Here we describe the design, fabrication and characterization of a planar, silicon-chip-based optomechanical crystal capable of co-localizing and strongly coupling 200-terahertz photons and 2-gigahertz phonons. These planar optomechanical crystals bring the powerful techniques of optics and photonic crystals to bear on phononic crystals, providing exquisitely sensitive (near quantum-limited), optical measurements of mechanical vibrations, while simultaneously providing strong nonlinear interactions for optics in a large and technologically relevant range of frequencies.

Many of the salient features of three-dimensional photonic crystal or photonic bandgap materials can be realized in lower-dimensional, planar waveguiding structures with the use of standard lithographic and etching procedures<sup>6</sup>. Here we study both the optical and mechanical properties of a textured silicon-on-insulator microchip, with the aim of demonstrating that photonic crystal structures can also be used to detect, generate, and control mechanical vibrations (phonons) within the same planar, chip-scale architecture.

The geometry of the optomechanical crystal structure considered here is shown in Fig. 1a. The effectively one-dimensional structure consists of a silicon nanobeam (thickness  $t$  not shown) with rectangular holes formed by thin cross-bars connected on both sides to thin rails. Figure 1b shows a finite-element-method (FEM) simulation of the optical band structure of such a periodically patterned nanobeam projection (we refer here to the infinitely periodic structure as a projection). The electric field profile for modes at the band edge ( $k_x = \pi/A$ , the boundary of the first Brillouin zone) are shown to the right of the band structure. The finite structure terminates at its supports on both ends, forming a doubly clamped beam. To form localized resonances in the centre of the structure, the discrete translational symmetry of the patterned beam is intentionally disrupted by a ‘defect’ consisting of a quadratic decrease in the lattice constant  $A$ , symmetric about the centre of the beam. The defect forms an effective potential for optical modes at the band edges, with the spatial dependence of the effective potential closely following the spatial properties of the defect<sup>17</sup> (see Fig. 1b inset). Thus the optical modes of the infinitely periodic structure are confined

by a quasi-harmonic potential. This effective potential localizes a ‘ladder’ of modes with Hermite–Gauss envelopes, analogous to the modes of the one-dimensional harmonic potential of quantum mechanics. The localized optical modes of the finite structure (hereafter referred to as Device 1) are also found by FEM simulation and shown in Fig. 1b to the right of the corresponding mode of the projection.

Analogously, Fig. 1c shows a FEM simulation of the mechanical band structure of the infinitely periodic nanobeam. Mechanical modes at the band edge experience an effective potential analogous to the optical modes, localizing certain types of vibrations to the defect region. Owing to symmetry (see Supplementary Information), only the coloured bands give rise to mechanical modes that, when localized by the defect, yield ladders of modes with strong dispersive coupling to the localized optical modes. We classify these optomechanically coupled mechanical modes, from lowest to highest frequency, as the ‘pinch’, ‘accordion’ and ‘breathing’ modes. Exemplary localized mechanical modes for each highlighted band are shown to the right of the corresponding mode of the projection.

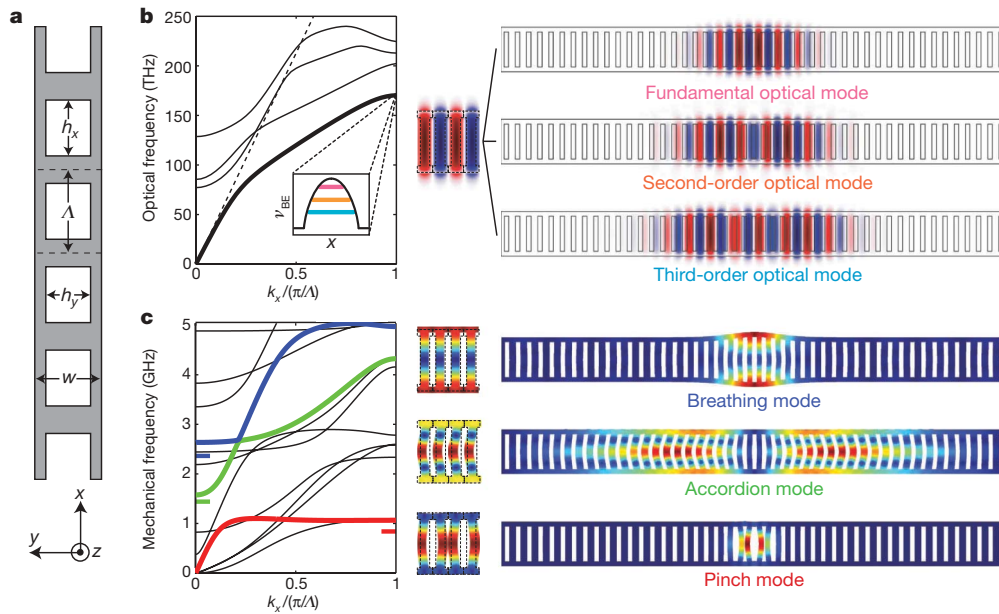
The two kinds of waves, mechanical and optical, are on equal footing in this structure. Each mechanical mode has a frequency  $\nu_m$  and displacement profile  $\mathbf{Q}(\mathbf{r})$ ; each optical mode has a frequency  $\nu_o$  and electric field profile  $\mathbf{E}(\mathbf{r})$ , where  $\mathbf{r}$  is the position vector. Just as the optical mode volume,  $V_o = \int dV (\sqrt{\epsilon} |\mathbf{E}| / \max(|\sqrt{\epsilon} \mathbf{E}|))^2$ , describes the electromagnetic localization of the optical mode within a material of dielectric constant  $\epsilon(\mathbf{r})$ , the mechanical mode volume,  $V_m \equiv \int dV (|\mathbf{Q}| / \max(|\mathbf{Q}|))^2$  (see Supplementary Information), describes the effective localization of the mechanical mode within a material of constant mass density and elastic moduli. For both the localized optical and mechanical modes of the patterned beam cavity, the effective mode volume is less than a cubic wavelength. The effective motional mass, being proportional to the mode volume ( $m_{\text{eff}} \equiv \rho V_m$ ; with silicon mass density  $\rho_{\text{Si}} = 2.33 \text{ g cm}^{-3}$ ), is between 50 femtograms (fg) and 1,000 fg for the mechanical modes shown in Fig. 1c.

Drawing on recent work in the field of cavity optomechanics<sup>18,19</sup>, we describe the coupling between optical and mechanical degrees of freedom (to lowest order) by an effective coupling length  $L_{\text{OM}} \equiv \left( \frac{1}{\nu_o} \frac{d\nu_o}{d\alpha} \right)^{-1}$ , where the dispersion  $d\nu_o/d\alpha$  is the rate of change in the frequency of an optical resonance caused by the mechanical displacement parameterized by  $\alpha$ . For this work,  $\alpha$  is defined as the maximum displacement of the mechanical mode pattern. By definition then, the smaller  $L_{\text{OM}}$ , the larger the optical response for a given mechanical displacement.  $L_{\text{OM}}$  is also the length over which a photon’s momentum is transferred into the mechanical mode as it propagates within the structure, and thus is inversely proportional to the force per photon applied to the mechanical system.

To calculate  $L_{\text{OM}}$ , we used a perturbative theory of Maxwell’s equations with respect to shifting material boundaries<sup>20</sup>. The derivative  $d\nu_o/d\alpha$  around some nominal position, where the optical fields are known, can be calculated exactly without actually deforming the structure for a surface-normal displacement of the boundaries,

<sup>1</sup>Thomas J. Watson Sr Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA.





**Figure 1 | Optomechanical crystal design.** **a**, Geometry of nanobeam structure. **b**, **c**, Optical (**b**) and mechanical (**c**) bands and modes calculated via FEM for the experimentally fabricated (Device 1) silicon nanobeam ( $A = 362$  nm,  $w = 1,396$  nm,  $h_y = 992$  nm,  $h_x = 190$  nm, and  $t = 220$  nm; isotropic Young's modulus of 168.5 GPa; refractive index  $n = 3.493$ ). The band diagrams and unit-cell modes are calculated for the infinitely periodic beam structure with the above parameters. The coloured bands in panel **c** correspond to optically coupled mechanical modes. The frequency of the fundamental localized mode for each coloured mechanical band is indicated

$h(\alpha; \mathbf{r}) \equiv \mathbf{Q}(\mathbf{r}) \cdot \hat{\mathbf{n}} = \alpha \mathbf{q}(\mathbf{r}) \cdot \hat{\mathbf{n}}$ , where  $\mathbf{q}(\mathbf{r}) = \mathbf{Q}(\mathbf{r})/\alpha = d\mathbf{Q}(\mathbf{r})/d\alpha$  is the unitless displacement profile of the mechanical mode, and  $\alpha$  parameterizes the amplitude of the displacement. Using this perturbative formulation of Maxwell's equations, we find:

$$\frac{1}{L_{\text{OM}}} = \frac{1}{2} \frac{\int dA \left( \frac{d\mathbf{Q}}{d\alpha} \cdot \hat{\mathbf{n}} \right) \left[ \Delta \varepsilon |\mathbf{E}_{\parallel}|^2 - \Delta(\varepsilon^{-1}) |\mathbf{D}_{\perp}|^2 \right]}{\int dV \varepsilon |\mathbf{E}|^2} \quad (1)$$

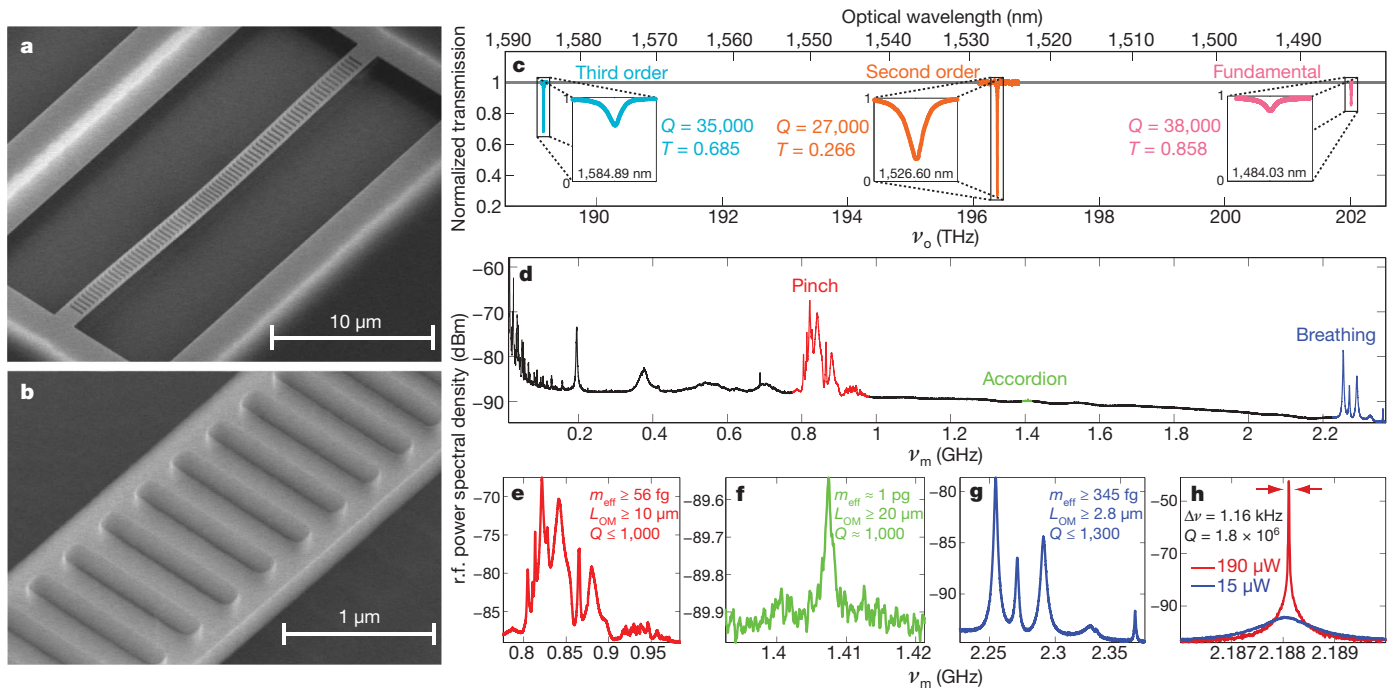
where  $\hat{\mathbf{n}}$  is the unit normal vector on the surface of the unperturbed cavity,  $\mathbf{D}(\mathbf{r}) = \varepsilon(\mathbf{r})\mathbf{E}(\mathbf{r})$ ,  $\Delta \varepsilon = \varepsilon_1 - \varepsilon_2$ ,  $\Delta(\varepsilon^{-1}) = \varepsilon_1^{-1} - \varepsilon_2^{-1}$ ,  $\varepsilon_1$  is the dielectric constant of the periodic structure, and  $\varepsilon_2$  is the dielectric constant of the surrounding medium ( $\varepsilon_2 = \varepsilon_0$  in this case). This method of calculating the coupling provides a wealth of intuition about the nature of the coupling and can be used to engineer the structure for strong optomechanical coupling.

Figure 2a and b shows scanning electron microscope images of a fabricated silicon nanobeam with the parameters of Device 1. The optical modes of the nanobeam are probed with a tapered and dimpled optical fibre<sup>21</sup> in the near-field of the defect cavity, simultaneously sourcing the cavity field and collecting the transmitted light in a single channel (see Methods). Figure 2c shows the low-pass filtered optical transmission spectrum of Device 1 at low optical input power ( $\sim 30$   $\mu\text{W}$ ). The optical cavity resonances are identified by comparison to FEM modelling of the optical modes of the structure. The radio frequency (r.f.) spectrum provides information about the optically coupled mechanical modes of the structure through the phase and amplitude modulation imparted to the transmitted light via the mechanical motion of these modes. Figure 2d–g shows the measured photodetector r.f. power spectral density of the optical transmission through the second-order optical cavity resonance (this mode was used because of its deep on-resonance coupling). A series of lower-frequency modes can be seen in the r.f. spectra ( $\sim 200$  MHz and higher harmonics), corresponding to compression modes of the entire beam, followed by groups of localized phononic modes of the lattice at 850 MHz (pinch), 1.41 GHz (accordion), and 2.25 GHz

by a horizontal bar of the same colour. The localized modes are calculated for a finite, patterned beam with a central defect (the inset in panel **b** shows the approximate shape of the optical band-edge frequency  $\nu_{\text{BE}}$  versus position  $x$  along the length of the beam due to the defect). The defect region consists of  $N_{\text{defect}} = 15$  central holes in which the hole-to-hole spacing is varied quadratically from the nominal lattice constant at the beam perimeter (362 nm) to 85% of that value for the two holes straddling the central cross-bar. The total number of holes in the finite beam structure is  $N_{\text{total}} = 75$ .

(breathing). The transduced signal at low optical power corresponds to thermally excited motion of the mechanical modes, and is inversely proportional to  $m_{\text{eff}} L_{\text{OM}}^2$  (see below). At higher optical input power (shown in Fig. 2h), regenerative mechanical oscillation<sup>22</sup> of the fundamental breathing mode is excited owing to the strong dynamical back-action<sup>18</sup> between the co-localized phonon and photon modes.

The scaling of frequency with the geometry of the observed vibrational modes can be used to differentiate the various types of modes that are coupled to the optical field. The top panel of Fig. 3a, for instance, shows the trend in frequency of the lowest-frequency mechanical mode of the manifold of modes around 2.3 GHz for a series of devices for which the in-plane geometry has been uniformly scaled. Identified as the fundamental breathing mode in Fig. 2g, the frequency of this mode scales nearly perfectly with the planar geometric scale factor, as expected for the predominantly in-plane vibrations of a thin beam structure. This should be contrasted with the optical modes (Fig. 3a, bottom panel), whose frequencies do not scale simply with the planar geometry, a result of the coupling of in-plane and vertical optical mode confinement. Figure 3b shows the result of a non-uniform in-plane scaling of the optomechanical crystal geometry. The r.f. power spectral density for Device 1 is plotted along with that of a second device, Device 2, which has a lattice constant,  $A = 365$  nm, and total length,  $L$ , that are essentially identical to those of Device 1, but has a considerably smaller width ( $w = 864$  nm,  $h_y = 575$  nm,  $h_x = 183$  nm). Simulations show that the pinch modes are the lowest-frequency group of localized and optically coupled mechanical modes in both structures (see right panels of Fig. 3b). From the calculated band structures, the ratio of the frequency of the localized pinch-mode manifold, after accounting for the defect, is  $1.826 \text{ GHz}/846 \text{ MHz} = 2.16$  for the two devices. Experimentally, the ratio of the fundamental pinch mode frequencies (highlighted in red) is  $1.749 \text{ GHz}/805 \text{ MHz} = 2.17$ . This should be compared with the second-order compression mode of the nanobeam, highlighted in yellow in Fig. 3b, which is the lowest-frequency mechanical mode



**Figure 2 | Photonic and phononic crystal mode spectroscopy.** **a, b**, Scanning electron microscope images of the fabricated silicon nanobeam optomechanical crystal. **c**, Optical spectroscopy of Device 1 with the taper waveguide in contact. The taper-loaded quality factor  $Q$  and on-resonance normalized transmission level  $T$  are indicated for each optical mode. **d**, Mechanical spectroscopy of Device 1 with taper waveguide in contact.

of the global beam structure with optomechanical coupling. This mode has a frequency that should be  $\frac{3\pi}{2L} \sqrt{\frac{E}{\langle \rho \rangle}}$ , where  $E$  is Young's modulus and  $\langle \rho \rangle$  is the average linear mass density. The frequency of this mode is measured to be 234 MHz/195 MHz = 1.20 times higher in Device 2 than in Device 1, which is in good agreement with the ratio  $\sqrt{\langle \rho_1 \rangle / \langle \rho_2 \rangle} = 1.23$ . The difference between the change in the frequencies of the lattice-localized versus beam modes illustrates the independence of these two 'systems'; once the wavelength of the global beam modes approaches the scale of the lattice periodicity, the vibrations become localized and behave independently of the global beam structure (such as the end clamps).

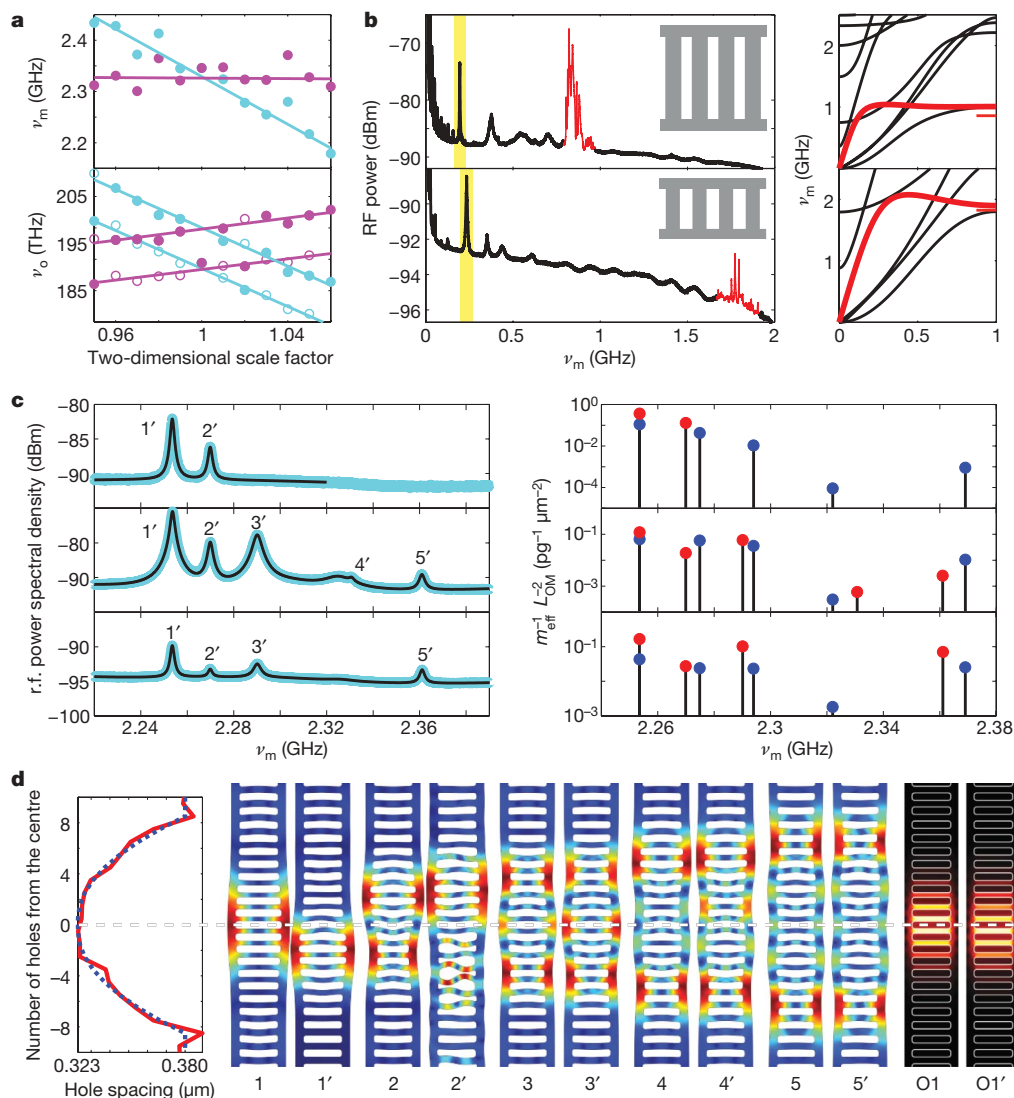
Further details of the properties of the lattice-localized phononic cavity modes can be determined through quantitative study of their optical transduction. Figure 3c shows the r.f. optical transmission spectrum (blue) and fit (black) due to Brownian motion of the breathing modes of Device 1 (that is, at low optical input power), for the three optical modes shown in Figs 1c and 2c. Because the various optical modes have different spatial profiles, each mechanical mode has a different  $L_{OM}$  for each optical mode. The root-mean-square (r.m.s.) mechanical amplitude of a mode due to Brownian motion is  $\langle \alpha^2 \rangle = k_B T / (m_{eff} Q^2)$ . It can be shown analytically that the factor  $1 / (m_{eff} L_{OM}^2)$  uniquely determines the transduction of the Brownian motion for these sideband-resolved optomechanical oscillations (see Supplementary Information). To the right of each measured spectrum is the experimentally extracted mechanical frequency and value of  $1 / (m_{eff} L_{OM}^2)$  (red), together with the values of these quantities obtained from equation (1) and the FEM model (blue). A good correspondence, in both frequency and transduced signal amplitude, is found across all optical and mechanical mode pairs. To achieve this level of correspondence, imperfections in the fabricated structure are taken into account by extracting the geometry from high-resolution scanning electron microscope images of the device and calculating the modified optical and mechanical modes (Fig. 3d). The resulting measured value for optomechanical

**e–g**, Zoomed r.f. mechanical spectra of Device 1 showing pinch (red), accordion (green) and breathing (blue) modes. **h**, r.f. spectra of the fundamental breathing mode (from a third device) at low (15 μW) and high (190 μW) coupled optical power. At high optical power the breathing mode is optically amplified through dynamical back-action of the radiation pressure force, with an above-threshold effective  $Q$ -factor greater than  $10^6$ .

coupling between the fundamental breathing and optical mode (assuming a FEM-calculated motional mass of  $m_{eff} = 330$  fg) is  $L_{OM} = 2.9$  μm, approaching the limit of half the effective in-material wavelength of light. The sensitivity of the mechanical transduction of the fundamental breathing mode can be appreciated by comparing the mode's r.m.s. thermal amplitude at  $T = 300$  K,  $\alpha_{thermal} = 245$  fm, to its quantum zero-point motion of  $\alpha_{zero-point} = 3.2$  fm. The sensitivity limit, as given by the background level in the middle panel of Fig. 3c, is thus a factor of  $\sim 7.5$  times that of the standard quantum limit.

The loss of mechanical energy from the confined mechanical modes of a phononic crystal can, in principle, be made arbitrarily low (and thus the mechanical  $Q$  arbitrarily high) by including a large number of unit cells outside the localizing defect region. Of course, other forms of mechanical damping, such as thermo-elastic damping, phonon–phonon scattering, or surface damping effects, would eventually become dominant<sup>23</sup>. For the fundamental breathing mode of the one-dimensional phononic crystal structures studied here, we measured an intrinsic mechanical  $Q$ -factor as high as 2,700 in air, at room temperature, and in contact with the taper waveguide (see the sub-threshold spectrum of Fig. 2h). This corresponds to a frequency- $Q$  product of  $\sim 6 \times 10^{12}$  Hz, a value close to the largest yet shown<sup>24</sup>. Although further tests (as a function of lattice periods, temperature, material processing, and so on) are required to determine the contribution of various mechanical loss mechanisms, numerical simulations show that mode coupling between localized and leaky phonon modes exist in these one-dimensional cavity structures and can significantly limit the  $Q$ -factor (Supplementary Information). This obstacle can be overcome in two-dimensional periodic slab structures, which have been shown to possess complete gaps for both optical and mechanical modes simultaneously<sup>25</sup>, indicating that optomechanical crystals may be used to probe sensitively the fundamental loss mechanisms within nanomechanical systems.

The experimental demonstration of optomechanical coupling between 200 THz photons and 2 GHz phonons in a planar optomechanical crystal points to new methods of probing, stimulating and



**Figure 3 | Phononic mode tuning and transduction.** **a**, Geometric scaling (planar) of the fundamental breathing mode. The frequencies plotted in magenta are the normalized frequencies, that is, the bare frequencies (cyan) times the scale factor for the device. The top panel shows the fundamental mechanical mode frequency, and the bottom panel shows the two lowest-order optical cavity mode frequencies. Device 1 is the device with scale factor 1.03. The best linear least-squares fit lines in the top panel correspond to a mechanical frequency change of  $-0.9\% \pm 0.2\%$  per device; the normalized mechanical frequency changes by  $-0.01\% \pm 0.2\%$  per device. The optical frequency of the mode used to make the mechanical measurement is filled (the other optical mode is open). **b**, Engineering of pinch mode frequencies, showing two devices (grey, insets) with pinch mode frequencies of 850 MHz (top) and 1.75 GHz (bottom). The mechanical band diagrams of each

using mechanical and optical interactions in a chip-scale platform. In particular, the efficient interconversion of photons and phonons enables a new class of optically powered oscillators<sup>22</sup> (such as that shown in Fig. 2h), phonon-mediated optical delay lines, and wide-band optical frequency conversion. The ability to generate, distribute, and interconnect optical and mechanical excitations on a single microchip enables more complex functionality, such as filters and mixers for r.f./microwave photonics signal processing<sup>26</sup>, non-reciprocal photonic structures for optical isolation<sup>27</sup>, and dynamic routing/buffering of optical signals<sup>28</sup>. The realization of optomechanical crystals in piezoelectric or light-emitting semiconductor materials such as gallium arsenide provides a further electronic interface for integration with microelectronic systems. In the rapidly developing field of mesoscopic quantum mechanical systems<sup>18,29</sup>, optomechanical crystals offer a unique architecture for such studies

structure are shown to the right of the measured r.f. spectrum, with the pinch mode band highlighted in red. **c**, Transduction of breathing mode motion. Left panels, measured r.f. spectrum of the optical transmission intensity for the first three optical modes (top to bottom panels), with breathing modes labelled from the fundamental mode (1') to the fifth order mode (5'). Right panels, corresponding measured (red) and modelled (blue) transduction amplitude and frequency of the breathing modes. See text and Supplementary Information for details. **d**, FEM-simulated ideal (unprimed) and actual (primed) breathing modes of the silicon nanobeam cavity due to the ideal (blue dashed curve) and actual (red solid curve) lattice pitch in the defect region of the cavity. The first five mechanical modes are shown, along with the fundamental optical mode 'O1'.

with their high frequency of operation and phononic bandgap shielding from the thermal bath. Finally, optomechanical crystals could also be used as high-spatial-resolution mass sensors: with  $m_{\text{eff}} = 62$  fg and  $\nu_m = 850$  MHz, the mass of a single haemoglobin A protein ( $\sim 10^{-19}$  g) would change the frequency of the pinch mode by 700 Hz, allowing a sensitivity similar to that of nano electro-mechanical systems zeptogram mass sensors<sup>30</sup>.

## METHODS SUMMARY

The optomechanical crystal nanobeam is formed in the 220-nm-thick silicon device layer of a [100] silicon-on-insulator wafer. The pattern is defined in electron beam resist by electron beam lithography. The resist pattern is transferred to the device layer by an inductively coupled plasma reactive ion etch with a  $\text{C}_4\text{F}_8/\text{SF}_6$  gas chemistry. The nanobeam is then undercut and released from the silica BOX (buried oxide) layer by wet undercutting with hydrofluoric acid.



The experimental set-up used to measure the optical, mechanical and optomechanical properties of the silicon optomechanical crystal consists of a bank of fibre-coupled tunable infrared lasers spanning approximately  $\Delta\lambda = 200$  nm in wavelength, centred around  $\lambda = 1,520$  nm. After passing through a variable optical attenuator and fibre polarization controller, light enters the tapered and dimpled optical fibre, the position of which can be controlled with nanometre-scale precision. The transmission from the fibre is (optionally) passed through another variable optical attenuator and finally reaches an avalanche photodiode with a transimpedance gain of  $1.1 \times 10^4$  and a bandwidth (3 dB rolloff point) of 1.2 GHz. The avalanche photodiode has an internal bias tee, and the r.f. voltage is connected to the 50-ohm input impedance of a high-speed oscilloscope. The oscilloscope can perform a Fourier transform to yield the r.f. power spectral density, which is calibrated using a frequency generator that outputs a variable-frequency sinusoid with known power.

Received 5 June; accepted 23 September 2009.

Published online 18 October 2009.

- Yablonovitch, E. Inhibited spontaneous emission in solid-state physics and electronics. *Phys. Rev. Lett.* **58**, 2059–2062 (1987).
- Cregan, R. F. *et al.* Single-mode photonic band gap guidance of light in air. *Science* **285**, 1537–1539 (1999).
- Notomi, M. *et al.* Extremely large group-velocity dispersion of line-defect waveguides in photonic crystal slabs. *Phys. Rev. Lett.* **87**, 253902 (2001).
- Akahan, Y., Asano, T., Song, B.-S. & Noda, S. Fine-tuned high-q photonic-crystal nanocavity. *Opt. Express* **13**, 1202–1214 (2005).
- Soljačić, M. *et al.* Photonic-crystal slow-light enhancement of nonlinear phase sensitivity. *J. Opt. Soc. Am. B* **19**, 2052–2059 (2002).
- McNab, S. J., Moll, N. & Vlasov, Y. A. Ultra-low loss photonic integrated circuit with membrane-type photonic crystal waveguides. *Opt. Express* **11**, 2927–2939 (2003).
- Olsson, R. H. III & El-Kady, I. Microfabricated phononic crystal devices and applications. *Meas. Sci. Technol.* **20**, 012002, 1–13 (2009).
- Kushwaha, M. S., Halevi, P., Dobrzynski, L. & Djafari-Rouhani, B. Acoustic band structure of periodic elastic composites. *Phys. Rev. Lett.* **71**, 2022–2025 (1993).
- Montero de Espinosa, F. R., Jiménez, E. & Torres, M. Ultrasonic band gap in a periodic two-dimensional composite. *Phys. Rev. Lett.* **80**, 1208–1211 (1998).
- Sánchez-Pérez, J. V. *et al.* Sound attenuation by a two-dimensional array of rigid cylinders. *Phys. Rev. Lett.* **80**, 5325–5328 (1998).
- Robertson, W. M. & Rudy, J. F. III. Measurement of acoustic stop bands in two-dimensional periodic scattering arrays. *J. Acoust. Soc. Am.* **104**, 694–699 (1998).
- Olsson, R. H. III, El-Kady, I. F., Su, M. F., Tuck, M. R. & Fleming, J. G. Microfabricated vhf acoustic crystals and waveguides. *Sens. Actuators A* **145–146**, 87–93 (2008).
- Khelif, A., Djafari-Rouhani, B., Vasseur, J. O. & Deymier, P. A. Transmission and dispersion relations of perfect and defect-containing waveguide structures in phononic band gap materials. *Phys. Rev. B* **68**, 024302 (2003).
- Trigo, M., Bruchhausen, A., Fainstein, A., Jusserand, B. & Thierry-Mieg, V. Confinement of acoustical vibrations in a semiconductor planar phonon cavity. *Phys. Rev. Lett.* **89**, 227402 (2002).
- Kang, M. S., Nazarkin, A., Brenn, A. & Russell, P. S. J. Tightly trapped acoustic phonons in photonic crystal fibres as highly nonlinear artificial raman oscillators. *Nature Phys.* **5**, 276–280 (2009).
- Maldovan, M. & Thomas, E. L. Simultaneous localization of photons and phonons in two-dimensional periodic structures. *Appl. Phys. Lett.* **88**, 251907 (2006).
- Chan, J., Eichenfield, M., Camacho, R. & Painter, O. Optical and mechanical design of a “zipper” photonic crystal optomechanical cavity. *Opt. Express* **17**, 3802–3817 (2009).
- Kippenberg, T. J. & Vahala, K. J. Cavity optomechanics: back-action at the mesoscale. *Science* **321**, 1172–1176 (2008).
- Favero, I. & Karrai, K. Optomechanics of deformable optical cavities. *Nature Phys.* **3**, 201–205 (2009).
- Johnson, S. G. *et al.* Perturbation theory for Maxwell’s equations with shifting material boundaries. *Phys. Rev. E* **65**, 066611 (2002).
- Michael, C. P., Borselli, M., Johnson, T. J., Chrystal, C. & Painter, O. An optical fiber-taper probe for wafer-scale microphotonic device characterization. *Opt. Express* **15**, 4745–4752 (2007).
- Rokhsari, H., Kippenberg, T. J., Carmon, T. & Vahala, K. J. Radiation-pressure-driven micromechanical oscillator. *Opt. Express* **13**, 5293–5301 (2005).
- Mohanty, P. *et al.* Intrinsic dissipation in high-frequency micromechanical resonators. *Phys. Rev. B* **66**, 085416 (2002).
- Weinstein, D. & Bhawe, S. A. Internal dielectric transduction of a 4.5 ghz silicon bar resonator. *IEEE Int. Electron. Devices Meeting* 415–418 (2007).
- Maldovan, M. & Thomas, E. Simultaneous complete elastic and electromagnetic band gaps in periodic structures. *Appl. Phys. B* **83**, 595–600 (2006).
- Hosseini-Zadeh, M. & Vahala, K. J. Photonic rf down-converter based on optomechanical oscillation. *IEEE Photon. Technol. Lett.* **20**, 234–236 (2008).
- Yu, Z. & Fan, S. Complete optical isolation created by indirect interband photonic transitions. *Nature Photon.* **3**, 91–94 (2009).
- Lin, Q. *et al.* Coherent mixing of mechanical excitations in nano-optomechanical structures. Preprint at <<http://arxiv.org/abs/0908.1128>> (2009).
- Schwab, K. C. & Roukes, M. L. Putting mechanics into quantum mechanics. *Phys. Today* **58**, 36–42 (2005).
- Yang, Y. T., Callegari, C., Feng, X. L., Ekinci, K. L. & Roukes, M. L. Zeptogram-scale nanomechanical mass sensing. *Nano Lett.* **6**, 583–586 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** Funding for this work was provided by a DARPA seed grant (grant no. HR0011-08-0002) and the National Science Foundation (EMT grant no. 0622246, MRSEC grant no. DMR-0520565, and CIAN grant no. EEC-0812072 through University of Arizona).

**Author Contributions** M.E., J.C., and R.C. performed the design, fabrication, and testing of devices. M.E., K.J.V., and O.P. developed the device concept and planned the measurements. All authors worked together to write the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to O.P. ([opainter@caltech.edu](mailto:opainter@caltech.edu)).

# Soft colloids make strong glasses

Johan Mattsson<sup>1†</sup>, Hans M. Wyss<sup>1†</sup>, Alberto Fernandez-Nieves<sup>1†</sup>, Kunimasa Miyazaki<sup>2†</sup>, Zhibing Hu<sup>3</sup>, David R. Reichman<sup>2</sup> & David A. Weitz<sup>1</sup>

Glass formation in colloidal suspensions has many of the hallmarks of glass formation in molecular materials<sup>1–5</sup>. For hard-sphere colloids, which interact only as a result of excluded volume, phase behaviour is controlled by volume fraction,  $\phi$ ; an increase in  $\phi$  drives the system towards its glassy state, analogously to a decrease in temperature,  $T$ , in molecular systems. When  $\phi$  increases above  $\phi^* \approx 0.53$ , the viscosity starts to increase significantly, and the system eventually moves out of equilibrium at the glass transition,  $\phi_g \approx 0.58$ , where particle crowding greatly restricts structural relaxation<sup>1–4</sup>. The large particle size makes it possible to study both structure and dynamics with light scattering<sup>1</sup> and imaging<sup>3,4</sup>; colloidal suspensions have therefore provided considerable insight into the glass transition. However, hard-sphere colloidal suspensions do not exhibit the same diversity of behaviour as molecular glasses. This is highlighted by the wide variation in behaviour observed for the viscosity or structural relaxation time,  $\tau_\alpha$ , when the glassy state is approached in supercooled molecular liquids<sup>5</sup>. This variation is characterized by the unifying concept of fragility<sup>5</sup>, which has spurred the search for a ‘universal’ description of dynamic arrest in glass-forming liquids. For ‘fragile’ liquids,  $\tau_\alpha$  is highly sensitive to changes in  $T$ , whereas non-fragile, or ‘strong’, liquids show a much lower  $T$  sensitivity. In contrast, hard-sphere colloidal suspensions are restricted to fragile behaviour, as determined by their  $\phi$  dependence<sup>1,6</sup>, ultimately limiting their utility in the study of the glass transition. Here we show that deformable colloidal particles, when studied through their concentration dependence at fixed temperature, do exhibit the same variation in fragility as that observed in the  $T$  dependence of molecular liquids at fixed volume. Their fragility is dictated by elastic properties on the scale of individual colloidal particles. Furthermore, we find an equivalent effect in molecular systems, where elasticity directly reflects fragility. Colloidal suspensions may thus provide new insight into glass formation in molecular systems.

The concept of fragility is best summarized in a renormalized Arrhenius plot, where the temperature is rescaled by the glass-transition temperature,  $T_g$ , and fragility is defined by the logarithmic slope at  $T_g$  (ref. 5). This representation highlights the wide variation in behaviour of molecular liquids, ranging from strong to fragile, and provides a unifying conceptual framework. An understanding of the physical origin of fragility, however, is lacking. The fragility of colloidal suspensions must be defined by their concentration dependence rather than by their  $T$  dependence. With this definition, colloidal hard-sphere suspensions are restricted to fragile behaviour and the absence of a wider range of fragilities limits their versatility as a model system of the glass transition.

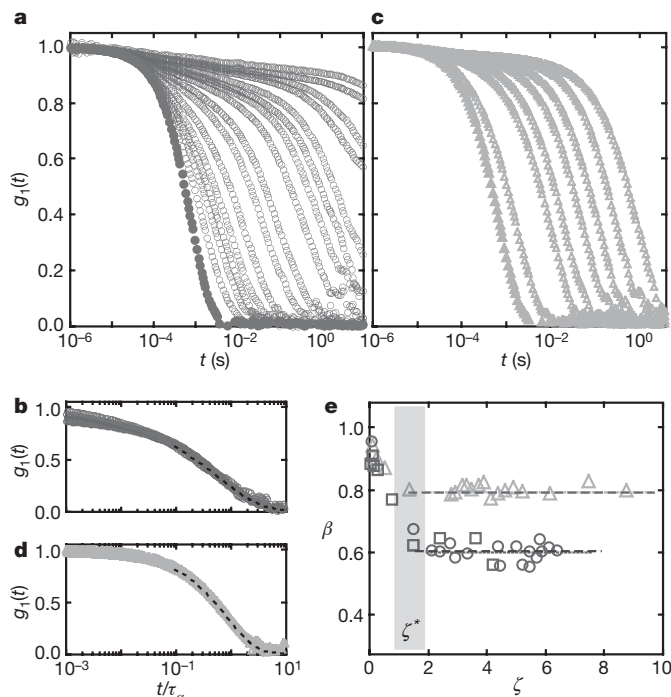
We studied aqueous suspensions of deformable microgel particles of varying elasticity (Methods Summary). To obtain a measure of

their elasticity, we probed the compressibility of individual particles by determining the change in particle size with both osmotic pressure and temperature (Supplementary Information). We determined the particle concentrations,  $\zeta$ , from the known polymer concentrations of the microgel suspensions. For hard-sphere suspensions, the particle concentration is quantified by the volume fraction,  $\phi = nV_p$ , where  $n$  is the number density of particles and  $V_p$  is their volume. However, because the microgel particles are deformable, their volume is not fixed and  $\phi$  is no longer a good measure of concentration. Instead, we used  $\zeta = nV_0$ , where  $V_0 = 4\pi R_0^3/3$  is the volume of an undeformed particle of radius  $R_0$  measured in dilute suspension;  $\zeta$  is always proportional to the polymer concentration. At low concentrations, where  $V_p$  is independent of concentration,  $\zeta = \phi$  and soft microgels can be treated as hard spheres<sup>7</sup>. To calibrate the measurement of  $\zeta$ , we determined the viscosity at low concentrations and used the Einstein expression to directly link the polymer concentration and  $\zeta$  (Supplementary Information). As  $\zeta$  increases above  $\phi_{RCP} \approx 0.63$ , corresponding to the random close packing of undeformed spheres, the particles must shrink; we confirmed this behaviour using static light scattering, which shows a systematic shift of the first peak in the effective structure factor (Supplementary Information).

We studied the dynamics of the microgel suspensions for fixed  $T$  and varying  $\zeta$  using dynamic light scattering (Supplementary Information). The behaviour observed for deformable microgel particles is qualitatively similar to that of hard spheres: for the lowest  $\zeta$  value, we find a monomodal decay of the intermediate scattering function,  $g_1(t)$ , which captures diffusion of particles in a dilute suspension, as shown for microgels with ‘intermediate’ stiffness and  $R_0 = 92$  nm in Fig. 1a. As  $\zeta$  increases, there are marked changes in the dynamics and we observe a two-step relaxation, characteristic of glass-forming materials upon approach to the glass transition<sup>5</sup>; the fast initial decay is only weakly dependent on  $\zeta$ , whereas the final decay, which reflects structural relaxation, depends strongly on  $\zeta$ . The final decay is well described by a stretched exponential,  $g_1(t) \propto \exp(-t/\tau)^\beta$ , where  $\tau$  is the relaxation time and  $\beta$  is the stretching exponent. Moreover, as for hard spheres, the final decay can be superimposed for different values of  $\zeta$  onto a single master curve by rescaling the decay times, as shown in Fig. 1b. We find that  $\beta = 0.6$ , as demonstrated in Fig. 1e. A corresponding superposition of the shape of the structural relaxation is often possible for molecular liquids near their glass-transition temperature, where it is referred to as time–temperature superposition<sup>5</sup>.

Unlike hard-sphere suspensions, the microgels are not dynamically arrested at  $\zeta \approx 0.63$ . Instead, it is necessary to consider much higher  $\zeta$  values, whereupon the final decay eventually exceeds any experimentally accessible timescale, ultimately resulting in formation of a glass, as shown in Fig. 1a. To quantify the  $\zeta$  dependence of the final decay, we determined the average structural relaxation time,  $\tau_\alpha = (\tau/\beta)\Gamma(1/\beta)$ ,

<sup>1</sup>Department of Physics and Harvard School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>2</sup>Department of Chemistry, Columbia University, New York, New York 10027, USA. <sup>3</sup>Department of Physics, University of North Texas, Denton, Texas 76203, USA. <sup>†</sup>Present addresses: Department of Applied Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden (J.M.); TU Eindhoven, Institute for Complex Molecular Systems and Department of Mechanical Engineering, PO Box 513, 5600 MB Eindhoven, The Netherlands (H.M.W.); School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA (A.F.-N.); Institute of Physics, University of Tsukuba, Tsukuba 305-8571, Japan (K.M.).



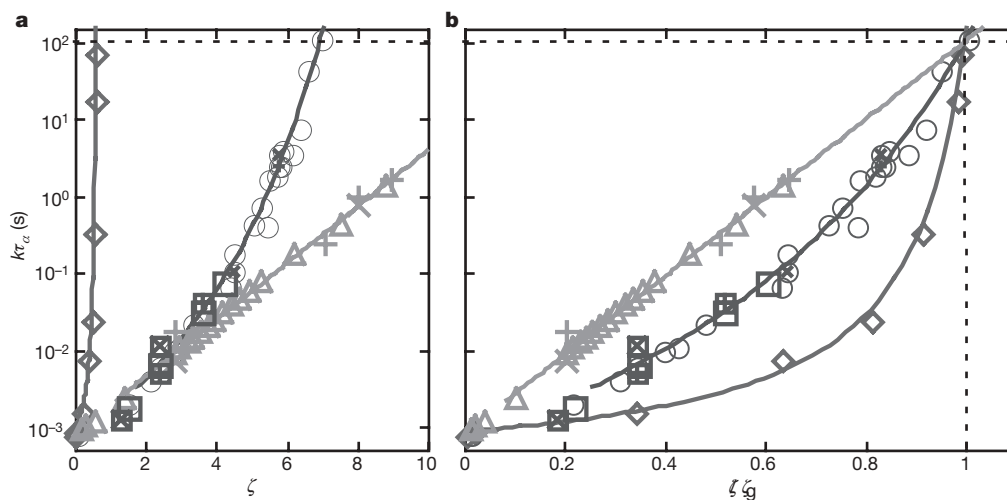
**Figure 1 | Dynamics of microgel suspensions.** **a, c,** Plots of  $g_1(t)$  for intermediate (**a**) and soft (**c**) microgels at  $Q = 23\mu\text{m}^{-1}$ , where  $\zeta$  ranges from  $8 \times 10^{-4}$  to 8.1 and from 0.15 to 9.4, respectively. Filled symbols indicate data for  $\zeta < \zeta^*$ . **b, d,** Time–concentration superposition for intermediate (**b**) and soft (**d**) microgels. The dashed lines are stretched exponential fits. **e,** The stretching exponent  $\beta$  remains constant for values of  $\zeta$  greater than  $\zeta^*$  (the range of  $\zeta^*$  is indicated by the shaded region) for both intermediate (circles,  $R_0 = 92$  nm; squares,  $R_0 = 168$  nm) and soft (triangles,  $R_0 = 80$  nm) microgels.

where  $\Gamma$  denotes the gamma function. We used a fixed value of the scattering vector,  $Q$ , for each sample even though the particle size,  $R$ , and hence  $QR$ , vary slightly with  $\zeta$ . In sharp contrast to the behaviour of hard spheres, where small changes in volume fraction lead to pronounced slowing of the structural relaxation, the dynamic arrest for microgel suspensions stretches over a wide range of  $\zeta$  values. This difference is highlighted by a comparison of the  $\zeta$  dependence of  $\tau_\alpha$  for a ‘stiff’ hard-sphere-like microgel<sup>8</sup> (Fig. 2a, diamonds) with that

for our intermediate soft microgels (Fig. 2a, circles and squares). We have rescaled the relaxation times by a constant,  $k$  (Supplementary Information), to compensate for the different diffusion timescales in the dilute limit and for the differences in  $QR_0$ .

For each microgel sample, we also determined the frequency-dependent shear moduli,  $G'(\omega)$  and  $G''(\omega)$ , by oscillatory rheology;  $G'$  has a plateau at high frequencies and a low-frequency structural relaxation (Supplementary Information), strongly reminiscent of the behaviour of concentrated hard-sphere suspensions<sup>2</sup>. We determined  $\tau_\alpha$  as the timescale corresponding to the frequency for which  $G' = G''$  and observed the same  $\zeta$  dependence of  $\tau_\alpha$  as was observed from light scattering, as shown by the plus symbols in Fig. 2a. Moreover, the viscosity of the microgel suspensions, determined at sufficiently low shear rates to preclude shear thinning, can also be directly scaled onto the relaxation-time data. Thus, the  $\zeta$  dependence characterizing structural relaxation is robust, as it is independently determined by several different methods. To describe the relaxation-time data, both for the intermediate and the stiff microgels, we used a Vogel–Fulcher–Tammann (VFT) function<sup>5</sup>,  $\tau_\alpha = \tau_0 \exp(A\zeta/(\zeta_0 - \zeta))$ , where  $\zeta_0$  sets the apparent divergence,  $A$  controls the growth of  $\tau_\alpha$  on approach to  $\zeta_0$  and  $\tau_0$  is the characteristic relaxation time for low values of  $\zeta$ . This empirical function, with  $1/\zeta$  exchanged for  $T$ , provides a good description of the  $T$  dependence of  $\tau_\alpha$  for supercooled molecular liquids near their glass transition<sup>5</sup>. These data thus further confirm that for suspensions of soft particles,  $\zeta$  plays a role analogous to that of  $1/T$  in molecular systems.

Our experiments show that for soft colloidal particles, the approach to the glass transition upon varying  $\zeta$  is more gradual than for hard spheres; this is reminiscent of the behaviour found in molecular liquids upon varying  $T$ , where different fragilities are observed. We therefore propose that the concept of fragility can be extended to colloidal systems through control of particle elasticity, provided that we use  $\zeta$  in the colloidal systems. To test this hypothesis, we investigated the behaviour of even softer microgels (Supplementary Information). These had similar behaviour in their dynamics, including a nearly exponential decay of  $g_1(t)$  at low  $\zeta$  values and a two-step decay with stretched-exponential behaviour of the final decay for  $\zeta$  values greater than a certain crossover concentration,  $\zeta^*$  (Supplementary Information), as shown in Fig. 1c. Moreover, the final decay can again be scaled onto a single master curve, as shown in Fig. 1d; however,  $\beta = 0.8$ , which is higher than for our microgel of intermediate stiffness (Fig. 1e). When  $\tau_\alpha$  is plotted as a function of  $\zeta$ , we find exponential behaviour,  $\tau_\alpha = \tau_0 \exp(C\zeta)$ , where



**Figure 2 | Fragility range for colloids.** **a,** Plot of  $k\tau_\alpha$  versus  $\zeta$  for stiff (diamonds,  $R_0 = 95$  nm)<sup>8</sup>, intermediate (empty circles,  $R_0 = 92$  nm) and soft (triangles,  $R_0 = 80$  nm) microgels, where  $k$  is chosen to collapse the data onto those of the intermediate sample at low  $\zeta$  values. Data for a second intermediate sample (empty squares,  $R_0 = 168$  nm) scale onto those of the

first for  $\zeta > \zeta^*$ , as expected. Rescaled shear viscosities (intermediate: crosses in circles,  $R_0 = 92$  nm, and crosses in squares,  $R_0 = 168$  nm; soft: crosses,  $R_0 = 80$  nm) and rheological structural relaxation times (intermediate: pluses in circles,  $R_0 = 92$  nm, and pluses in squares,  $R_0 = 168$  nm; soft: pluses,  $R_0 = 80$  nm). **b,** Same as **a**, with  $\zeta$  normalized by  $\zeta_g = \zeta(\tau_\alpha = 100$  s).



$C$  is a constant, as shown by the triangles in Fig. 2a; the same behaviour is observed in measurements of both rheology (pluses) and viscosity (crosses). This exponential dependence is analogous to the Arrhenius behaviour typically observed in strong molecular glass-formers such as silica, but here with  $T$  exchanged for  $1/\zeta$ ; the exponential dependence also corresponds to VFT behaviour with  $\zeta_0 \gg \zeta$ .

We thus confirm our hypothesis that the concept of fragility can be extended to colloidal systems by control of the particle elasticity and through the variation of  $\zeta$ . Interestingly, we also find the smallest stretching exponent for the sample with the strongest dynamic behaviour. This observation is reminiscent of that seen in molecular systems, where a correlation between higher  $\beta$  values and stronger dynamic behaviour has been reported<sup>5</sup>.

To explore the analogy between soft colloidal suspensions and molecular glass-formers further, we represent our data in a rescaled Arrhenius plot in a fashion similar to that used for molecular glasses<sup>5</sup>. For molecular glass-formers, the abscissa of this plot is typically scaled by  $T_g$ , defined as the temperature ( $\tau_\alpha = 100$  s) (ref. 5), where the structural relaxation is no longer experimentally accessible. By analogy, we rescale the concentration  $\zeta$  by  $\zeta_g = \zeta(\tau_\alpha = 100$  s) to obtain the corresponding plot for colloids. We view the slope of the data at  $\zeta_g$  as the fragility. As in molecular liquids, the fragility varies continuously from strong to fragile, as shown in Fig. 2b; for the microgels, it is controlled by particle elasticity. We note that the ratio  $\zeta_g/\zeta^*$  is smaller the more fragile is the behaviour, just as the ratio  $T^*/T_g$  is smaller the more fragile is a molecular liquid<sup>5</sup>; here,  $T^*$  denotes the crossover temperature (Supplementary Information). Changes in molecular motions, characteristic of a deeply supercooled liquid, are first observed at  $T^*$ , corresponding to  $\zeta^*$  in colloids. We note that pressure-dependent measurements on molecular liquids can be summarized by a similar type of plot of viscosity as a function of inverse volume rescaled by the glass-transition volume; this suggests a connection between volume-dependent fragility in colloids and molecular liquids<sup>9</sup>.

To investigate the role of particle elasticity in the dynamic arrest, we used oscillatory rheology to determine the concentration dependence of the elastic modulus on a timescale characterizing local particle motion,  $G_p(\zeta)$  (Methods Summary). We determined the elastic energy on a particle length scale  $G_p(\zeta)V_p$ , where  $V_p$ , the volume of the deformable particles, is estimated for  $\zeta \geq \zeta^*$  to be  $V_p = (\zeta^*/\zeta)4\pi R_0^3/3$ . This elastic energy dominates over thermal energy near the glass transition for all samples, as demonstrated by the dependence of  $G_p(\zeta)V_p/k_B T$  on  $\zeta/\zeta^*$  (Fig. 3a;  $k_B$ , Boltzmann constant); here we scaled by  $\zeta^*$  to mark the onset of the relevant dynamic range. Moreover, we observe a direct correlation between the growth of this energy ratio and fragility: the softer the particles, the stronger their suspension behaviour and, correspondingly, the weaker the growth of elastic energy with  $\zeta$  as it increases above  $\zeta^*$ . This suggests that the

suspension elasticity on the particle scale controls the dynamics above  $\zeta^*$  and, thus, determines the fragility. To compare data for different fragilities, we determined the volumes  $V^* = \gamma V_p$  (Supplementary Information), where  $\gamma$  is a constant chosen to make the elastic energy equal the thermal energy at  $\zeta^*$ . The energy ratios increase more quickly as the fragility increases, as shown by a plot of  $G_p(\zeta)V^*/k_B T$  as a function of  $\zeta/\zeta^*$  (Fig. 3b). We find that  $V^*$  is smaller than the particle volume for all samples, confirming that the elasticity on local length scales determines the fragility.

For supercooled liquids, there are apparent correlations between fragility and the behaviour of the elastic properties<sup>10–15</sup> as expressed, for example, by the non-ergodicity parameter<sup>12</sup>, the bulk and shear moduli<sup>10,11,13–15</sup> or the strength of the boson peak<sup>14</sup>; however, the underlying origin of fragility in supercooled liquids remains unresolved. On the basis of our results for deformable colloids, we explored the possibility that the elastic energy correlates with fragility also for molecular systems. We determined the elastic energy,  $G_\infty(T)V_{liq}^*$ , where  $G_\infty(T)$  denotes the high-frequency shear modulus as measured by Brillouin light scattering and  $V_{liq}^*$ , the volume for each liquid, is chosen such that  $G_\infty(T^*)V_{liq}^* = k_B T^*$ . We find that  $G_\infty(T)V_{liq}^*/k_B T$  as a function of  $T^*/T$  is well described by a linear function for each liquid; moreover, an increase in fragility is accompanied by a more pronounced slope, as shown in Fig. 3c. In addition,  $V_{liq}^*$  is smaller than the molecular size for all liquids (Supplementary Information), in further analogy with our findings for soft, deformable colloids. The striking qualitative similarity in the behaviour of the elastic energies of molecular systems and colloidal suspensions confirms the important role of elasticity in determining fragility. Our results thus suggest that thermal activation controls fragility, with a  $\zeta$ -dependent elastic energy for colloids, corresponding to a  $T$ -dependent elastic energy for molecular liquids.

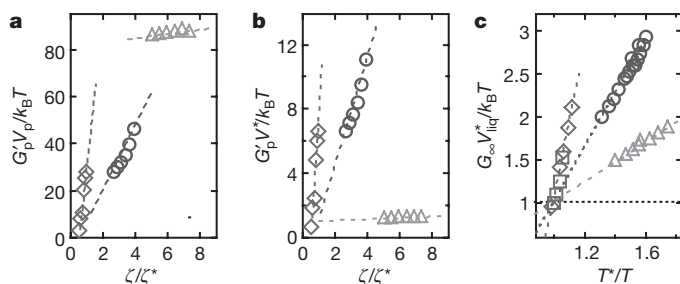
Structural relaxation in molecular glass-formers can be described by thermal activation within a complex energy landscape, which reflects the possible configurations of the system; fragility is reflected in the detailed topology of this energy landscape<sup>5,16</sup>. Fragility might also have a parallel in the temperature dependence of purely thermodynamic properties, such as entropy<sup>17,18</sup>. Interestingly, theoretical work for colloidal hard-sphere systems suggests that the  $\phi$  variation of the local configurational properties is surprisingly similar to the  $T$  variation of the corresponding configurational properties of molecular liquids<sup>19</sup>. This implies that the concept of fragility is far more general than previously believed. Our work provides more evidence for this by demonstrating that the concept of fragility can be directly extended to suspensions of deformable colloidal particles, where particle concentration controls glass formation. Furthermore, our work suggests that the origin of the variation in fragility resides in the elastic properties of the particles themselves; soft particles lead to strong behaviour and hard particles lead to fragile behaviour.

In molecular glass-formers, strong dynamic behaviour generally reflects a significant degree of directional bonding that is typical of network-forming glass-formers such as silica. By analogy, we speculate that the origin of the behaviour of colloidal suspensions is similar: softer particles are more easily deformed, and as the particles shrink their shapes will increasingly deviate from the spherical, which in turn leads to an increased directionality of interparticle interactions<sup>20,21</sup>, corresponding to strong behaviour. By contrast, hard particles retain their spherical shape and their interactions are more isotropic, corresponding to fragile behaviour.

The results presented here demonstrate the remarkable similarities between the behaviours of colloidal and molecular glass-formers, and suggest that an improved understanding of glass formation may come from exploring the properties of other soft colloidal systems, such as solutions of star polymers<sup>22</sup> and block-copolymer micelles<sup>23</sup>.

## METHODS SUMMARY

**Microgel properties.** Our microgel particles consisted of interpenetrated and crosslinked polymer networks of poly(*N*-isopropylacrylamide) and polyacrylic acid<sup>24</sup>. To study the glass-forming properties of the microgel suspensions, we



**Figure 3 | Elasticity of glass-formers.** **a**, Data for soft (triangles) and intermediate (circles) microgel suspensions compared with data for a suspension of model hard spheres (diamonds)<sup>25</sup>. For the hard-sphere system,  $\zeta^* = \phi^* = 0.53$  (ref. 26). The lines are guides to the eye. **b**, A different representation of the data in **a**. **c**, Data for orthoterphenyl (diamonds)<sup>27</sup>, salol (squares)<sup>28</sup>, glycerol (circles)<sup>29</sup> and silica (triangles)<sup>30</sup>. The dashed lines are fits to the data. The plot is explained in detail in the text.

used systems with size polydispersities of  $\sim 40\%$ , which successfully suppressed crystallization. In addition, our samples were naturally well matched in both density and refractive index, as the particles contained a large fraction of water. The synthesis parameters sensitively affected the elastic properties of the microgel particles themselves. We determined the elastic properties by investigating the change in particle size with osmotic pressure or with temperature (Supplementary Information).

**Elasticity of the microgel suspensions.** The viscoelastic shear moduli  $G'(\omega)$  and  $G''(\omega)$ , measured for different suspension concentrations, can be superimposed on a master curve by scaling the frequency. Moreover, the master curves for different suspensions can themselves be superimposed onto one common master curve by scaling the magnitudes of the moduli (Supplementary Information). This behaviour enabled us to determine the elastic response at a frequency corresponding to local particle motion,  $\omega_p$ , determined as the characteristic rheological timescale at  $\xi^*$  (Supplementary Information). For the intermediate and the soft samples, we determined  $G'_p = G'(\omega_p)$ , the elastic modulus at  $\omega_p$ , and from  $G'_p$  we determined the effective elastic energy per particle,  $G'_p V_p$ . We compared our results with data for a hard-sphere-like system from ref. 25.

Received 30 April; accepted 21 August 2009.

- van Meegen, W. & Pusey, P. N. Dynamic light-scattering study of the glass transition in a colloidal suspension. *Phys. Rev. A* **43**, 5429–5441 (1991).
- Mason, T. G. & Weitz, D. A. Linear viscoelasticity of colloidal hard sphere suspensions near the glass transition. *Phys. Rev. Lett.* **75**, 2770–2773 (1995).
- Kegel, W. K. & van Blaaderen, A. Direct observation of dynamical heterogeneities in colloidal hard-sphere suspensions. *Science* **287**, 290–293 (2000).
- Weeks, E. R., Crocker, J. C., Levitt, A. C., Schofield, A. & Weitz, D. A. Three-dimensional direct imaging of structural relaxation near the colloidal glass transition. *Science* **287**, 627–631 (2000).
- Angell, C. A., Ngai, K. L., McKenna, G. B., McMillan, P. F. & Martin, S. W. Relaxation in glass-forming liquids and amorphous solids. *J. Appl. Phys.* **88**, 3113–3157 (2000).
- Cheng, Z., Zhu, J., Chaikin, P. M., Phan, S.-E. & Russel, W. B. Nature of the divergence in low shear viscosity of colloidal hard-sphere dispersions. *Phys. Rev. E* **65**, 041405 (2002).
- Stieger, M., Pedersen, J. S., Lindner, P. & Richtering, W. Are thermoresponsive microgels a model system for concentrated colloidal suspensions? A rheology and small-angle neutron scattering study. *Langmuir* **20**, 7283–7292 (2004).
- Bartsch, E., Antonietti, M., Schupp, W. & Sillescu, H. The glass transition dynamics of polymer microneutral colloids. A mode coupling analysis. *J. Chem. Phys.* **97**, 3950–3963 (1992).
- Herbst, C. A., Cook, R. L. & King, H. E. Jr. Density-mediated transport and the glass transition: high pressure viscosity measurements in the diamond anvil cell. *J. Non-Cryst. Solids* **172–174**, 265–271 (1994).
- Tobolsky, A. V., Powell, R. E. & Eyring, H. in *Frontiers in Chemistry* Vol. 1 (eds Burk, R. E. & Grummit, O.) 125–190 (Interscience, 1943).
- Nemilov, S. V. *Thermodynamic and Kinetic Aspects of the Vitreous State* (CRC Press, 1995).
- Scopigno, T., Ruocco, G., Sette, F. & Monaco, G. Is the fragility of a liquid embedded in the properties of its glass? *Science* **302**, 849–852 (2003).
- Novikov, V. N. & Sokolov, A. P. Poisson's ratio and the fragility of glass-forming liquids. *Nature* **431**, 961–963 (2004).
- Novikov, V. N., Ding, Y. & Sokolov, A. P. Correlation of fragility of supercooled liquids with elastic properties of glasses. *Phys. Rev. E* **71**, 061501 (2005).
- Dyre, J. C. The glass transition and elastic models of glass-forming liquids. *Rev. Mod. Phys.* **78**, 953–972 (2006).
- Sastry, S., Debenedetti, P. G. & Stillinger, F. H. Signatures of distinct dynamical regimes in the energy landscape of a glass-forming liquid. *Nature* **393**, 554–557 (1998).
- Ito, K., Moynihan, C. T. & Angell, C. A. Thermodynamic determination of fragility in liquids and a fragile-to-strong liquid transition in water. *Nature* **398**, 492–495 (1999).
- Martinez, L.-M. & Angell, C. A. A thermodynamic connection to the fragility of glass-forming liquids. *Nature* **410**, 663–667 (2001).
- Bramer, Y. & Reichman, D. R. Mean-field theory, mode-coupling theory, and the onset temperature in supercooled liquids. *Phys. Rev. E* **69**, 041202 (2004).
- Wyart, M. On the rigidity of amorphous solids. *Ann. Phys. Fr.* **30**, 1–96 (2005).
- Shintani, H. & Tanaka, H. Frustration on the way to crystallization in glass. *Nature Mater.* **7**, 870–877 (2008).
- Roovers, J. Concentration dependence of the relative viscosity of star polymers. *Macromolecules* **27**, 5359–5364 (1994).
- Buitenhuis, J. & Förster, S. Block copolymer micelles: viscoelasticity and interaction potential of soft spheres. *J. Chem. Phys.* **107**, 262–272 (1997).
- Xia, X. & Hu, Z. Synthesis and light scattering study of microgels with interpenetrating polymer networks. *Langmuir* **20**, 2094–2098 (2004).
- Shikata, T. & Pearson, D. S. Viscoelastic behaviour of concentrated spherical suspensions. *J. Rheol.* **38**, 601–616 (1994).
- Kumar, S., Szamel, G. & Douglas, J. F. Nature of the breakdown in the Stokes-Einstein relationship in a hard sphere fluid. *J. Chem. Phys.* **124**, 214501 (2006).
- Wang, C. H., Zhu, X. R. & Shen, J. C. Angular and temperature dependence of depolarized and polarized Rayleigh-Brillouin spectra of a supercooled liquid: o-terphenyl. *Mol. Phys.* **62**, 749–764 (1987).
- Enright, G. D. & Stoicheff, B. P. Light scattering from shear modes in liquid salol. *J. Chem. Phys.* **64**, 3658–3665 (1976).
- Scarponi, F., Comez, L., Fioretto, D. & Palmieri, L. Brillouin light scattering from transverse and longitudinal acoustic waves in glycerol. *Phys. Rev. B* **70**, 054203 (2004).
- Polian, A., Vo-Thanh, D. & Richet, P. Elastic properties of a-SiO<sub>2</sub> up to 2300 K from Brillouin scattering measurements. *Europhys. Lett.* **57**, 375–381 (2002).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We are grateful to J. Zhou for technical assistance, to J. Bergholtz for complimentary use of his light-scattering equipment and to J.-W. Kim for discussions. This work was supported by the US National Science Foundation and Harvard University's Materials Research Science and Engineering Center; by the Hans Werthén Foundation, the Wenner-Gren Foundation, the Knut and Alice Wallenberg Foundation and the Royal Society of Arts and Sciences in Göteborg (J.M.); by the Ministerio de Ciencia e Innovación and the University of Almería (A.F.-N.); and by KAKENHI (K.M.).

**Author Contributions** J.M. and H.M.W. designed the study, performed the experiments, analysed and interpreted the data and wrote the manuscript; A.F.-N. designed the study, interpreted the data and wrote the manuscript; Z.H. designed the samples and contributed to the writing of the manuscript; K.M., D.R.R. and D.A.W. each contributed to data interpretation and the writing of the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to J.M. ([johanm@chalmers.se](mailto:johanm@chalmers.se)).

# Long aftershock sequences within continents and implications for earthquake hazard assessment

Seth Stein<sup>1</sup> & Mian Liu<sup>2</sup>

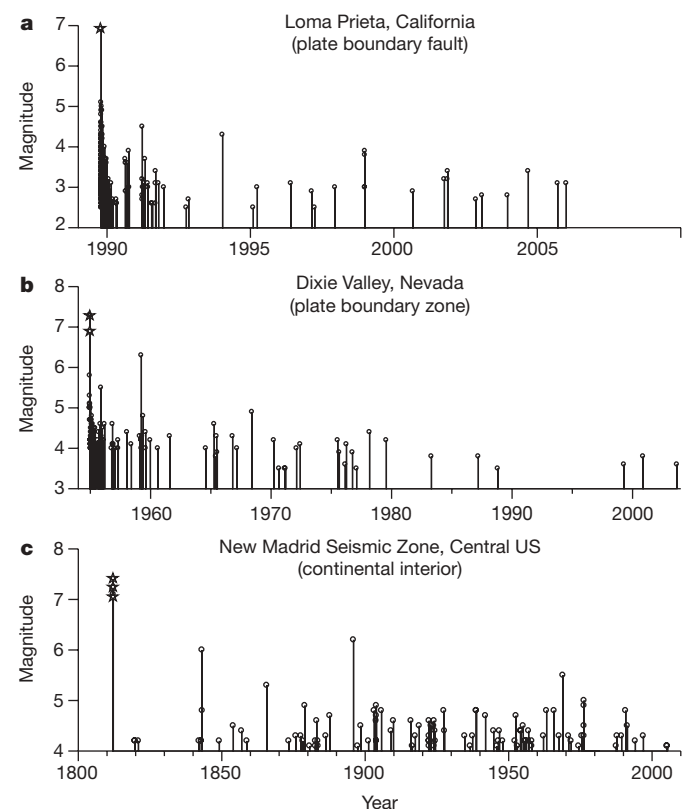
One of the most powerful features of plate tectonics is that the known plate motions give insight into both the locations and average recurrence interval of future large earthquakes on plate boundaries. Plate tectonics gives no insight, however, into where and when earthquakes will occur within plates, because the interiors of ideal plates should not deform. As a result, within plate interiors, assessments of earthquake hazards rely heavily on the assumption that the locations of small earthquakes shown by the short historical record reflect continuing deformation that will cause future large earthquakes<sup>1</sup>. Here, however, we show that many of these recent earthquakes are probably aftershocks of large earthquakes that occurred hundreds of years ago. We present a simple model predicting that the length of aftershock sequences varies inversely with the rate at which faults are loaded. Aftershock sequences within the slowly deforming continents are predicted to be significantly longer than the decade typically observed at rapidly loaded plate boundaries. These predictions are in accord with observations. So the common practice of treating continental earthquakes as steady-state seismicity overestimates the hazard in presently active areas and underestimates it elsewhere.

The disastrous magnitude-7.9 earthquake in Sichuan, China, in May 2008 was a surprise because it occurred on a fault that had had little recent seismicity. Such surprises occur often for earthquakes within continents. In contrast to plate boundaries where large (magnitude  $M > 7$ ) earthquakes occur at expected locations along the boundary faults, continental interiors like Sichuan or the central and eastern United States contain many old faults, most of which show little seismicity over the past hundred years for which we have seismological data. Hence, we are uncertain of the times and locations of future large earthquakes. Present earthquake hazard assessments typically assume that recent small earthquakes indicate the location of large future earthquakes. But what if some of these recent earthquakes are instead aftershocks of earlier large events?

Large earthquakes are typically followed by aftershock activity that decays to a lower level interpreted as 'normal' background seismicity<sup>2</sup>. This transition is difficult to identify precisely, because defining it depends on the area treated as the aftershock zone and the criterion used. For the majority of large earthquakes, which occur at plate boundaries, the transition generally occurs less than a decade after the main shock, as shown by both studies of individual events<sup>3–5</sup> and a large global compilation<sup>6</sup> (Fig. 1a). This duration is therefore regarded as the norm. In contrast, aftershock sequences can be much longer in other tectonic settings. In diffuse plate boundary zones such as the Basin and Range, aftershocks often continue for fifty years or more<sup>7–10</sup> (Fig. 1b). Within continental interiors, aftershocks may continue for hundreds of years after the main shock. Seismicity in the areas of past large earthquakes, including those in New Madrid, Missouri (1811–1812), Charlevoix, Quebec (1663), and Basel,

Switzerland (1356), may be aftershocks<sup>11</sup>. In the New Madrid seismic zone, the best-studied of these areas, seismicity delineates the areas thought to have ruptured in the main shocks, appears to be decreasing (Fig. 1c) and the largest earthquakes are at the ends of the presumed ruptures; all three observations are often found in aftershock studies<sup>12</sup>.

For most aftershock sequences, the number of earthquakes decreases approximately hyperbolically with time after the main-shock, following the empirical Omori's Law. Hence the aftershock durations may be estimated, using the change in seismicity rates from decaying aftershocks to background seismicity, which can be directly measured for short aftershock durations<sup>2,5,6</sup>. However, identifying this



**Figure 1 | Aftershock sequences following large earthquakes in different tectonic settings.** **a**, The Loma Prieta earthquake ( $M = 6.9$ ) on the San Andreas Fault on 17 October 1989. **b**, The Dixie Valley and Fairview Peak earthquakes ( $M = 7.1$  and  $M = 6.8$ ) in Nevada in December 1954. **c**, The New Madrid earthquakes ( $M = 7.0$ – $7.4$ ) in the central USA, that occurred December 1811–February 1812. Note the different time and magnitude scales for each panel. Stars are mainshocks. Data sources are given in Supplementary Information.

<sup>1</sup>Department of Earth and Planetary Sciences, Northwestern University, Evanston, Illinois 60208, USA. <sup>2</sup>Department of Geological Sciences, University of Missouri, Columbia, Missouri 65211, USA.



transition becomes difficult and then impossible as the aftershock duration approaches and then exceeds the length of the short earthquake record (Fig. 1). For historic earthquakes, further problems arise from incomplete and nonuniform catalogues. Nonetheless, bounds on the duration of some aftershock sequences can be inferred. A lower bound comes from treating today's earthquakes as aftershocks, and an upper bound comes where there is paleoseismic evidence of large (probably magnitude greater than 7) earthquakes but little seismicity today.

Figure 2 shows aftershock durations for selected events from three tectonic settings: plate boundary faults, broad plate boundary zones, and continental interiors. Despite the uncertainties in estimating both the duration and the rate at which tectonic slip loads the faults, the data indicate an inverse relationship between the aftershock durations and the slip rates. Faults at plate boundaries that are loaded by the rapid (typically faster than  $10 \text{ mm yr}^{-1}$ ) plate motion show aftershock durations of about ten years<sup>3–6</sup>. Faults within broad plate boundary zones but off the main boundaries move at only a small fraction of the plate motion (a few millimetres per year<sup>7–9</sup>) and have longer aftershock durations. Aftershocks continue today following the 1952 Kern County, California, and 1959 Hebgen Lake, Montana, earthquakes, and have a typical duration of about 100 years in the Central Nevada seismic belt<sup>10</sup>. Such a duration is consistent with the absence of aftershocks from large earthquakes on the Wasatch fault, the most recent of which occurred about 600 years ago<sup>9</sup>.

The longest aftershock sequences occur within continental plate interiors, which deform at rates typically less than a millimetre per year<sup>13–15</sup>. Today's New Madrid and Charlevoix seismicity give lower duration bounds. Upper bounds come from the absence of seismicity at sites of past earthquakes such as the Reelfoot Rift's eastern margin<sup>16</sup> in the New Madrid area and the Meers fault in Oklahoma<sup>17</sup>. Thus, within continents, aftershock sequences can last hundreds of years or longer.

It is reasonable that within continents aftershocks should continue long after large earthquakes. Aftershocks result from changes of stress and fault properties induced by the main shock. At a plate boundary, steady plate motion quickly reloads the fault after a large earthquake and overwhelms the effects of the main shock. Within continents, however, the faults are reloaded much more slowly, allowing aftershocks to continue much longer.

This inverse correlation between aftershock durations and fault loading rates is consistent with the rate-and-state model of fault

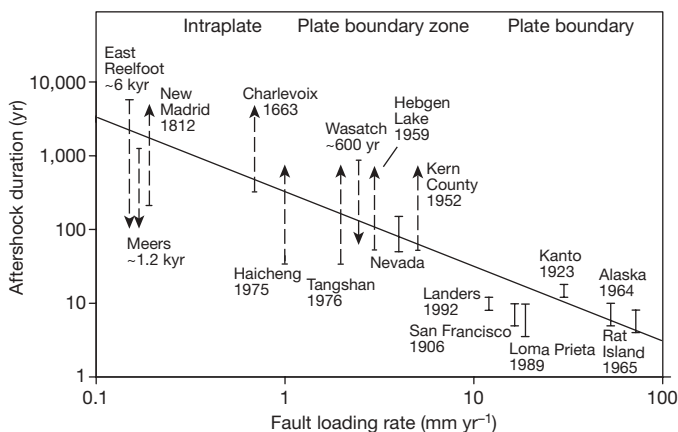
friction, which predicts changes in fault properties after earthquakes, and is commonly used for aftershock studies<sup>3–6,18</sup>. This model predicts an aftershock duration  $t_a = A\sigma/\dot{\tau}$ , where  $\dot{\tau}$  is the rate of shear stressing on the fault,  $\sigma$  is the normal stress, and  $A$  is a constitutive parameter<sup>3</sup>. Although the stressing rate is hard to measure, it can be estimated from the relative velocity (loading rate) across the fault. For a simple geometry<sup>19</sup> in which  $\dot{\tau} = \mu v/\pi w$ , where  $\mu$  is the rigidity,  $v$  is the loading rate, and  $w$  is the vertical extent of the seismogenic fault,  $t_a = A\sigma\pi w/\mu v$ . Using  $A = 0.01$ ,  $\sigma = 15 \text{ MPa}$ ,  $\mu = 30 \text{ GPa}$  and  $w = 20 \text{ km}$ , we can predict  $t_a = 314/v$  for  $t_a$  in years and  $v$  in millimetres per year. This  $t_a$ – $v$  relation is generally consistent with observations spanning a wide range of loading rates and aftershock durations (Fig. 2).

Both the data and the model used here are simplifications of complicated systems. The aftershock durations are inferred by several techniques, and include both aftershocks on the fault plane itself and ones in the surrounding area triggered by stress transfer<sup>4,6,20</sup>. The model treats all faults as having the same geometry and physical properties, and differing only in loading rate. Hence the single line shown represents a trend about which we expect scatter.

Although long aftershock sequences within continents are more easily identifiable because of low background seismicity, the systematic increase in aftershock durations for lower loading rates indicates that this variation reflects the physics of faulting. This variation can be explained by the rate-and-state model that attributes aftershocks to changes in fault friction after the main shock. Alternatively, aftershocks have been interpreted as resulting from stress disturbances. One such process is viscous relaxation in creeping fault segments surrounding the rupture zone<sup>21</sup>. This process occurs primarily within the upper crust for a few years after the main shock. However, large earthquakes can significantly load the viscous lower crust and upper mantle, which then transfer stress back to the seismogenic upper crust over a much longer time, as indicated by geodetic data<sup>22</sup>. This process would last longer within continents than near plate boundaries because of the more viscous lower crust and upper mantle, and so could also contribute to long aftershock sequences. The resulting exponential decay in seismicity would be similar to that predicted by Omori's law.

Recognizing such long aftershock sequences is crucial for seismic hazard assessment within continents. The observation that the locations of small continental earthquakes tend to be the sites of many future ones<sup>23</sup> is consistent with many of these events being part of long aftershock sequences. In this case, the locations of these small earthquakes may not indicate the timing or the locations of future large earthquakes, which are often episodic, temporally clustered, and migrate between faults over thousands of years<sup>13,24–28</sup>. A striking example of this variability is in North China, where large ( $M > 7$ ) earthquakes have been frequent, but not a single pair has occurred in the same place since 1300 AD. Several effects presumably contribute to these complex spatiotemporal patterns. Temporal clustering of earthquakes arises in various tectonic settings owing to stress transfer caused by earthquake interactions<sup>29</sup>. The slow tectonic loading within continents causes stress transfer to have a greater effect than at rapidly loaded plate boundaries<sup>27,28</sup>. The spatial migration of large earthquakes is probably related to fault interaction over multiple time-scales<sup>27</sup>. Because the tectonic loading is slow and is accommodated by a network of faults, factors such as the heterogeneity of lithospheric rheology, fault properties and geometry, and the earthquake history of each fault may all contribute to the complex spatiotemporal patterns.

The complex spatiotemporal patterns of large earthquakes, and the long durations of their aftershocks, make assessing earthquake hazards within continents difficult. In the short term much of the seismic hazard results from aftershocks, which can be damaging, and thus should not be removed in attempts to infer earthquake recurrence and hazards. In the longer term, relying unduly on recent seismicity to predict the locations of future large earthquakes will overestimate the hazard in some places and lead to surprises elsewhere.



**Figure 2 | Aftershock durations and fault loading rates for selected large earthquakes in different tectonic settings.** The solid line is the model prediction. Aftershock durations have bars showing the span of published estimates. One-sided constraints have bars at the known value and arrows indicating the open value. For example, the New Madrid aftershocks span at least 200 years, but the upper bound is unconstrained. Similarly, the Wasatch aftershocks span less than 600 years, but the lower bound is unconstrained. Data and sources for this plot are in the Supplementary Information. 1 kyr = 1,000 years.

Improved assessments of earthquake hazard require treating the networks of faults within continents as complex systems and developing a better understanding of how such systems behave in time and space<sup>28</sup>. This is being done by combining the seismological record with palaeoseismic studies that provide a longer earthquake history, and high-precision Global Positioning System (GPS) measurements that show whether deformation continues. Taken together, these data can show whether recent seismic events in a region are aftershocks or indicate future large earthquakes. One example is the New Madrid area, where palaeoseismic data show a cluster of large earthquakes in the past millennium<sup>25,30</sup> but two decades of GPS measurements show no deformation, so the cluster of large earthquakes there appears to have ended<sup>13,14</sup>.

Received 13 May; accepted 11 September 2009.

- Shedlock, K. M., Giardini, D., Grünthal, G. & Zhang, P. The GSHAP global seismic hazard map. *Seismol. Res. Lett.* **71**, 679–689 (2000).
- Ogata, Y. & Shimazaki, K. Transition from aftershock to normal activity: the 1965 Rat Islands earthquake aftershock sequence. *Bull. Seismol. Soc. Am.* **74**, 1757–1765 (1984).
- Dieterich, J. H. A constitutive law for rate of earthquake production and its application to earthquake clustering. *J. Geophys. Res.* **99**, 2601–2618 (1994).
- Toda, S., Stein, R. S., Reasenberg, P. A., Dieterich, J. H. & Yoshida, A. Stress transferred by the 1995 Mw=6.9 Kobe, Japan shock: effect of aftershocks and future earthquake probabilities. *J. Geophys. Res.* **103**, 24,543–24,565 (1998).
- Ziv, A. Does aftershock duration scale with mainshock size? *Geophys. Res. Lett.* **33**, doi:10.1029/2006GL027141 (2006).
- Parsons, T. Global Omori-law decay of triggered earthquakes: large aftershocks outside the classical aftershock zone. *J. Geophys. Res.* **107**, doi:10.1029/2001JB000646 (2002).
- Hammond, W. C., Kreemer, C. & Blewitt, G. in *Late Cenozoic Structure and Evolution of the Great Basin—Sierra Nevada Transition* 34–54 (Geol. Soc. Am. Spec. Pap. 447, 2009).
- Wesnowsky, S. G. Earthquakes, Quaternary faults, and seismic hazard in California. *J. Geophys. Res.* **91**, 12587–12631 (1986).
- Chang, W. L. & Smith, R. B. Integrated seismic-hazard analysis of the Wasatch front, Utah. *Bull. Seismol. Soc. Am.* **92**, 1904–1922 (2002).
- Ryall, A. Earthquake hazard in the Nevada region. *Bull. Seismol. Soc. Am.* **67**, 517–532 (1977).
- Ebel, J. E., Bonjer, K. P. & Oncescu, M. C. Paleoseismicity: seismicity evidence for past large earthquakes. *Seismol. Res. Lett.* **71**, 283–294 (2000).
- Stein, S. & Newman, A. Characteristic and uncharacteristic earthquakes as possible artifacts: applications to the New Madrid and Wabash seismic zones. *Seismol. Res. Lett.* **75**, 170–184 (2004).
- Newman, A. *et al.* Slow deformation and lower seismic hazard at the New Madrid Seismic Zone. *Science* **284**, 619–621 (1999).
- Calais, E. & Stein, S. Time-variable deformation in the New Madrid seismic zone. *Science* **5920**, 1442 (2009).
- Mazzotti, S., James, T., Henton, J. & Adams, J. GPS crustal strain, postglacial rebound, and seismic hazard in eastern North America: the Saint Lawrence valley example. *J. Geophys. Res.* **110**, doi:10.1029/2004JB0035900 (2005).
- Tuttle, M. P., Al-Shukri, H. & Mahdi, H. Very large earthquakes centered southwest of the New Madrid seismic zone 5,000–7,000 years ago. *Seismol. Res. Lett.* **77**, 361–380 (2006).
- Crone, A. J. & Luza, K. V. Style and timing of Holocene surface faulting on the Meers fault, southwestern Oklahoma. *Geol. Soc. Am. Bull.* **102**, 1–17 (1990).
- Toda, S. & Stein, R. Response of the San Andreas fault to the 1983 Coalinga-Nunez earthquakes: an application of interaction-based probabilities for Parkfield. *J. Geophys. Res.* **107**, doi:10.1029/2001JB000172 (2002).
- Savage, J. C. & Burford, R. O. Geodetic determination of relative plate motion in central California. *J. Geophys. Res.* **78**, 832–845 (1973).
- Stein, R. The role of stress transfer in earthquake occurrence. *Nature* **402**, 605–609 (1999).
- Zoller, G., Hainzl, S., Holschneider, M. & Ben-Zion, Y. Aftershocks resulting from creeping sections in a heterogeneous fault. *Geophys. Res. Lett.* **32**, doi:10.1029/2004GL021871 (2005).
- Hearn, E. H., Bürgmann, R. & Reilinger, R. E. Dynamics of Izmit earthquake postseismic deformation and loading of the Düzce earthquake hypocenter. *Bull. Seismol. Soc. Am.* **92**, 172–193 (2002).
- Kafka, A. in *Continental Intraplate Earthquakes: Science, Hazard, and Policy Issues* (eds Stein, S. & Mazzotti, S.) 35–48 (Geol. Soc. Am. Spec. Pap. 425, 2007).
- Crone, A. J. *et al.* Paleoseismicity of two historically quiescent faults in Australia: implications for fault behavior in stable continental regions. *Bull. Seismol. Soc. Am.* **93**, 1913–1934 (2003).
- Holbrook, J. *et al.* Stratigraphic evidence for millennial-scale temporal clustering of earthquakes on a continental-interior fault: Holocene Mississippi River floodplain deposits, New Madrid seismic zone, USA. *Tectonophysics* **420**, 431–454 (2006).
- Camelbeeck, T. *et al.* in *Continental Intraplate Earthquakes: Science, Hazard, and Policy Issues* (eds Stein, S. & Mazzotti, S.) 193–224 (Geol. Soc. Am. Spec. Pap. 425, 2007).
- Li, Q., Liu, M., & Stein, S. Spatiotemporal complexity of continental intraplate seismicity: insights from geodynamic modeling and implications for seismic hazard estimation. *Bull. Seismol. Soc. Am.* **99**, 52–60 (2009).
- Stein, S., Liu, M., Calais, E. & Li, Q. Mid-continent earthquakes as a complex system. *Seismol. Res. Lett.* **80**, 551–553 (2009).
- Kagan, Y. Y. & Jackson, D. D. Worldwide doublets of large shallow earthquakes. *Bull. Seismol. Soc. Am.* **89**, 1147–1155 (1999).
- Tuttle, M. P. *et al.* The earthquake potential of the New Madrid seismic zone. *Bull. Seismol. Soc. Am.* **92**, 2080–2089 (2002).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank Y. Yang for assistance in preparing the figures and Supplementary Information, and H. Wang for providing the Chinese earthquake catalogue. M.L.'s research is supported by NSF/OISE grant 0730154 and a grant from the Chinese Academy of Sciences. We also thank E. Calais, A. Friedrich, R. Smith and R. Stein for discussions, and T. Parsons and Y. Kagan for reviews.

**Author Contributions** S.S. developed the friction-based model and much of the data set. M.L. developed the Chinese data and the viscous relaxation model. Both authors discussed the results and participated in the writing.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to S.S. ([seth@earth.northwestern.edu](mailto:seth@earth.northwestern.edu)).

## LETTERS

## Experimental evolution of bet hedging

Hubertus J. E. Beaumont<sup>1,2,†</sup>, Jenna Gallie<sup>1</sup>, Christian Kost<sup>1,†</sup>, Gayle C. Ferguson<sup>1</sup> & Paul B. Rainey<sup>1</sup>

**Bet hedging—stochastic switching between phenotypic states<sup>1–3</sup>—is a canonical example of an evolutionary adaptation that facilitates persistence in the face of fluctuating environmental conditions. Although bet hedging is found in organisms ranging from bacteria to humans<sup>4–10</sup>, direct evidence for an adaptive origin of this behaviour is lacking<sup>11</sup>. Here we report the *de novo* evolution of bet hedging in experimental bacterial populations. Bacteria were subjected to an environment that continually favoured new phenotypic states. Initially, our regime drove the successive evolution of novel phenotypes by mutation and selection; however, in two (of 12) replicates this trend was broken by the evolution of bet-hedging genotypes that persisted because of rapid stochastic phenotype switching. Genome re-sequencing of one of these switching types revealed nine mutations that distinguished it from the ancestor. The final mutation was both necessary and sufficient for rapid phenotype switching; nonetheless, the evolution of bet hedging was contingent upon earlier mutations that altered the relative fitness effect of the final mutation. These findings capture the adaptive evolution of bet hedging in the simplest of organisms, and suggest that risk-spreading strategies may have been among the earliest evolutionary solutions to life in fluctuating environments.**

Life exists in ever-changing environments, but surviving under fluctuating conditions poses challenges. One solution is the evolution of mechanisms that allow modulation of phenotype in response to specific environmental cues. An alternative solution is stochastic phenotype switching, a strategy based on bet hedging, rather than direct environmental sensing<sup>4–10</sup>.

The general prediction from theory is that fluctuating selection generated by unpredictable environments can favour the evolution of bet hedging<sup>1–4,12–14</sup>. Under such conditions, a strategy that generates random variation in fitness-related traits among individuals within a population can enhance long-term fitness by increasing the likelihood that a subset of individuals expresses a phenotype that will be adaptive in a future environment<sup>3,15,16</sup>. However, the outcome of adaptive evolution under fluctuating selection is also shaped by factors such as the frequency of environmental change<sup>12</sup>, the capacity of a given population to respond to fluctuations by mutation and selection (that is, evolvability), the presence of suitable environmental cues<sup>13</sup> and the cost–benefit balance of different strategies<sup>12</sup>.

Here we report the *de novo* evolution of bet hedging in experimental bacterial populations. Our populations experienced repeated bouts of selection in two contrasting environments; they also experienced fluctuating selection wrought by imposition of an exclusion rule and population bottleneck. Applied at the point of transfer between environments, the exclusion rule assigned a fitness of zero to the type that was common in the current environment; imposition of the bottleneck meant that only a single phenotypically distinct type was selected from among the survivors to found the next bout of selection. The exclusion rule imposed strong selection for phenotypic innovation, whereas the

bottleneck negated the cost of bet hedging—that is, the generation of types maladapted to the prevailing conditions—by eliminating competition with conspecifics. A natural analogue of this mode of fluctuating selection is imposed by the host immune system on invading microorganisms. Indeed, many pathogens have evolved bet-hedging strategies based on stochastic antigen switching<sup>8</sup>.

When the ancestral genotype of *Pseudomonas fluorescens* SBW25 is grown in static broth microcosms, it rapidly diversifies into a range of niche specialist genotypes by mutation and selection, which each form distinct colonies on agar plates<sup>17</sup>. In contrast, diversification is constrained in shaken microcosms, which favour genotypes with an ancestral colony type<sup>18</sup>. We exploited the pleiotropic correlation between niche specialization and colony morphology to realize our selection regime. The expected evolutionary response was repeated evolution, fixation and extinction of genotypes with novel colony morphologies.

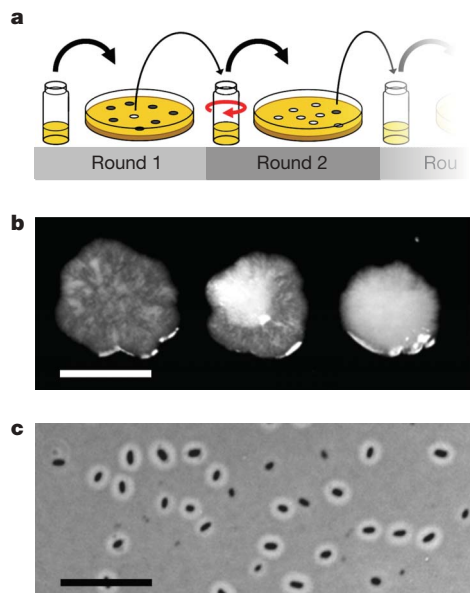
Twelve replicate selection lines were founded with the ancestral genotype and subjected to 16 rounds of alternating selection in static and shaken microcosms. During each round, populations were propagated by serial dilution until the emergence of cells that formed colonies with a heritable morphology different from that of their immediate ancestor. Selection for different colonies was open-ended rather than for an *a priori* defined morphology. Once detected, cells derived from a single individual of this new type were transferred to the opposing environment for the next round of selection (Fig. 1a and Methods). Thus, we imposed selection for a high growth rate in static and shaken microcosms, and simultaneously fluctuating selection for colony innovation. Each lineage was associated with a cognate control line that was under stabilizing selection for the ancestral colony morphology, but otherwise treated identically (see Methods).

This selection regime indeed drove the repeated evolution of new colony morphologies. However, here we concentrate on the evolution of genotypes capable of rapid, stochastic colony-morphology switching (hereafter referred to as colony switching), which emerged in two of 12 replicates (1B<sup>4</sup> and 6B<sup>4</sup>, Supplementary Note 1). Each of these switching genotypes formed distinct translucent and opaque colonies (Fig. 1b). Colony switching persisted in both lines for seven additional rounds of selection, after which the experiment was ended. None of the control lines gave rise to colony switching.

To obtain insight into the mechanisms by which switching types persisted, we re-imposed our selection regime on the genotype 1B<sup>4</sup> (Fig. 1b), but this time ignored the colony variants it generated through colony switching. Without exception, 1B<sup>4</sup> was driven to extinction by the evolution of new genotypes with novel colony morphologies, some of which did not switch (Supplementary Note 2). This shows that persistence was attributable to bet hedging—that is, the capacity to generate at high-frequency colonies that were phenotypically distinct from those selected in the previous round—rather than to an intrinsic growth rate advantage over non-switching types. This finding also draws attention to the importance of the bottleneck: in its absence,

<sup>1</sup>New Zealand Institute for Advanced Study and Allan Wilson Centre for Molecular Ecology & Evolution, Massey University, Private Bag 102904, North Shore Mail Centre, North Shore City 0745, Auckland, New Zealand. <sup>2</sup>Institute of Biology Leiden, Leiden University, PO Box 9505, 2300 RA Leiden, The Netherlands. <sup>†</sup>Present addresses: Institute of Biology Leiden, Leiden University, PO Box 9505, 2300 RA Leiden, The Netherlands (H.J.E.B.); Department of Bioorganic Chemistry, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany (C.K.).

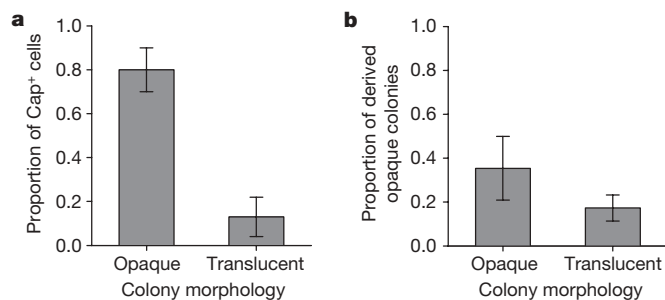




**Figure 1 | Colony morphology evolution.** **a**, Populations were propagated in static or shaken (red arrow) microcosms, and periodically screened for cells that form novel colony types, a prerequisite for passage to the opposing environment (the exclusion rule). Propagation between environments occurred by transfer of cells taken from a single, numerically dominant, novel colony (the bottleneck). **b**, From left: translucent, sectored and opaque colonies of a bet-hedging genotype ( $1B^4$ ). Sectored colonies, which resulted from stochastic capsule-expression switching during formation of the colony, were counted as opaque. Scale bar, approximately 2 mm. **c**, Capsulated and non-capsulated cells of  $1B^4$  (phase-contrast light microscopy with negative capsule staining). The proportions of capsulated cells in colonies produced by the original ancestor ( $1A^0$ ) and immediate ancestor ( $1A^4$ ) of  $1B^4$  were three orders of magnitude lower ( $n = 5$ , 500 cells examined per colony, binomial 95% confidence intervals 0.0010–0.0037 and 0.0028–0.0057, respectively). Scale bar, approximately 10  $\mu$ m.

$1B^4$  would have been eliminated by non-switching genotypes with faster growth rates.

The serendipitous discovery that centrifugation of  $1B^4$  cells resulted in two discrete fractions prompted microscopic examination (Supplementary Method 2), which revealed that  $1B^4$  produced both capsulated ( $Cap^+$ ) and non-capsulated ( $Cap^-$ ) cells (Fig. 1c). Opaque colonies contained a higher proportion of  $Cap^+$  cells than translucent colonies (Fig. 2a). When cell suspensions were plated from either colony type, both gave rise to a mixture of opaque and translucent colonies; however, cell suspensions made from opaque



**Figure 2 | Link between colony and cell morphology.** **a**, Mean proportion of capsulated cells in opaque and translucent colonies of bet-hedging genotype  $1B^4$  ( $n = 6$ , 100 cells examined per colony). Opaque colonies contained proportionally more capsulated cells than translucent colonies (analysis of deviance,  $F(1,10) = 61.795$ ,  $P < 0.0001$ ). Error bars, one standard deviation. **b**, Mean proportion of opaque colonies formed by cell suspensions prepared from the opaque and translucent colonies in **a** ( $n = 6$ , between 200 and 700 colonies examined). Cells derived from opaque colonies formed proportionally more opaque colonies (analysis of deviance,  $F(1,10) = 11.836$ ,  $P = 0.0063$ ). Error bars, one standard deviation.

colonies produced a higher proportion of opaque colonies than those derived from translucent colonies (Fig. 2b). Together, this suggests that the founding cell of a colony may determine the morphology of the latter by biasing the ratio of  $Cap^+$  and  $Cap^-$  cells. This causal link was corroborated by transposon mutagenesis<sup>19</sup>, which also indicated that the capsules consist of colanic acid<sup>20</sup>, a previously described capsule polysaccharide (see Supplementary Method 1 and Supplementary Table 1). Colony switching in the second bet-hedging genotype also involved  $Cap^+$  and  $Cap^-$  cells.

A likely mechanism of colony switching is reversible on–off switching of capsule production. Alternatively, colony switching may involve distinct non-switching  $Cap^+$  and  $Cap^-$  genotypes that form mixed colonies upon plating owing to chance co-localization, or mixed cell-type aggregates. We distinguished between these possibilities by determining whether populations that were passed through two sequential single-cell bottlenecks produced both  $Cap^+$  and  $Cap^-$  cells after each bottleneck (see Supplementary Method 2 and Supplementary Fig. 1). All replicate populations generated both cell types after each bottleneck. The probability of observing this pattern under the null hypotheses of no reversible switching is exceedingly low (one-tailed Fisher's exact test,  $P = 0.0004$ ; see Supplementary Note 3), indicating that  $1B^4$  switches reversibly between  $Cap^+$  and  $Cap^-$ .

Although bet hedging facilitated the long-term persistence of  $1B^4$ , the evolutionary emergence of this genotype required it to reach a detectable frequency within a selection round. To examine if  $1B^4$  owed its emergence to a higher fitness than its immediate ancestor ( $1A^4$ ), we competed these two genotypes in static microcosms, the environment in which  $1B^4$  emerged. This indicated that  $1B^4$  was indeed more fit than  $1A^4$  (one-sample  $t$ -test,  $n = 8$ ,  $P = 0.0002$ ; Fig. 4).

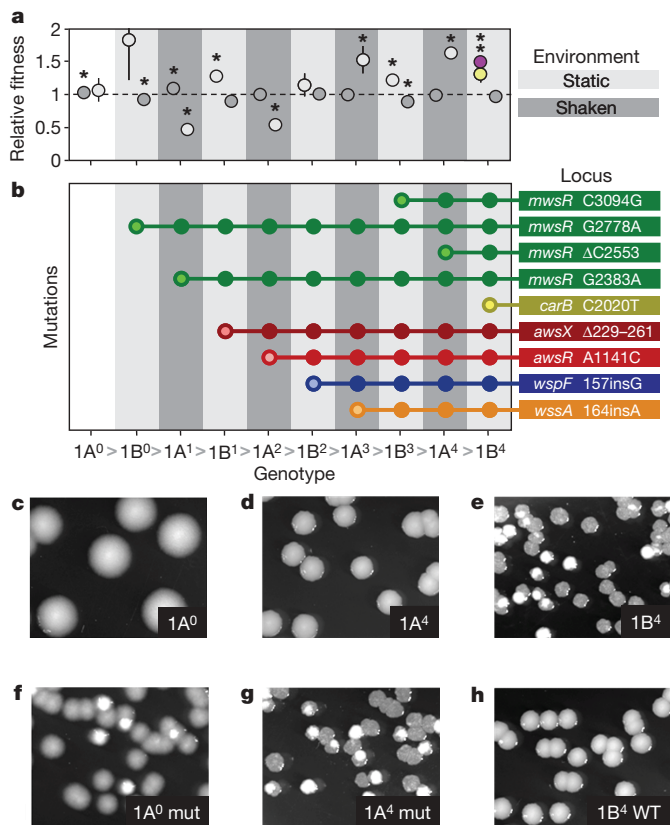
Repeated single-cell bottlenecks can drive the fixation of random deleterious mutations, causing a decline in fitness. To test if this had occurred during our experiment, we measured the fitness of all genotypes in the  $1B^4$  lineage relative to the original ancestor in both static and shaken microcosms (Fig. 3a). This revealed no evidence for a decrease in fitness. Interestingly, the results indicate non-transitive fitness interactions between some consecutive genotypes (for example,  $1A^4$  is more fit than  $1B^4$  relative to  $1A^0$ , but less fit than  $1B^4$  during direct competition).

Using whole-genome re-sequencing<sup>21,22</sup>, the entire 6.7-megabase-pair genome of  $1B^4$  (ref. 23) was analysed to unravel its mutational history. Nine mutations separating  $1B^4$  from the original ancestor were identified, confirmed by Sanger sequencing, and ordered by inspection of the affected loci in the preceding genotypes (Fig. 3b). With the exception of the final mutation, all mutations involved non-synonymous changes at loci previously demonstrated to be mutational targets in the evolution of wrinkly spreader types<sup>18,24–26</sup> (Fig. 3b). The final mutation was a single non-synonymous nucleotide change in *carB* (Arg674Cys, Fig. 3b), which encodes the large subunit of carbamoylphosphate synthetase (CarAB, EC 6.3.5.5), a central enzyme of the pyrimidine and arginine biosynthetic pathways<sup>27</sup>.

To examine the causal connection between the *carB* mutation and colony switching, we introduced this mutation in the immediate ancestor ( $1A^4$ ) by allelic replacement. The engineered genotype displayed colony switching (Fig. 3g). Conversely, reversion of the *carB* mutation in  $1B^4$  to wild type abolished colony switching (Fig. 3h). This demonstrates that the *carB* mutation is sufficient and necessary to cause stochastic colony morphology switching in  $1A^4$ .

Although the exact mechanism of colony switching remains to be elucidated, two lines of evidence suggest that switching might be controlled epigenetically, rather than by a mutable locus<sup>8</sup>. First,  $Cap^+$  and  $Cap^-$  cells were identical at the *carB* locus. Second, neither transposon mutagenesis nor genome re-sequencing showed evidence of the involvement of mutable loci (Supplementary Method 1).

If the *carB* mutation is the sole cause of the evolution of colony switching from  $1A^4$ , it must confer not only colony switching but also the requisite high fitness in static microcosms in this genetic background. To assess this, we introduced the mutant *carB* allele in  $1A^4$

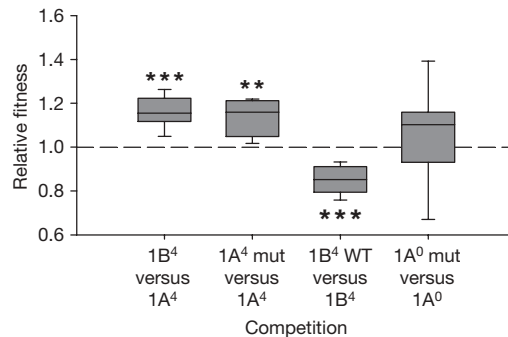


**Figure 3 | Fitness, mutational history and effects of the *carB* mutation in a bet-hedging lineage.** **a**, Fitness of each genotype in a bet-hedging lineage ( $1B^4$ ) measured relative to the original ancestor ( $1A^0$ ) in shaken (dark circles) and static (light circles) microcosms. Light bars, genotypes that evolved in static microcosms ( $1B^n$ ); dark bars, genotypes that evolved in shaken microcosms ( $1A^n$ ). Fitness of  $1B^4$  in static microcosms founded with predominantly capsulated (yellow circles,  $Cap^+$  proportion = 0.688, 95% confidence interval 0.417–0.872) or non-capsulated cells (purple circles,  $Cap^+$  proportion = 0.013, 95% confidence interval 0.004–0.031) did not differ significantly. Asterisks, significant deviation from 1 (dashed line;  $P \leq 0.05$ , one-sample *t*-tests,  $n = 3$ ). Error bars, one standard deviation. **b**, Mutations identified by genome re-sequencing of  $1B^4$  and ordered through analysis of the preceding genotypes. Each round of selection caused a genetic change readily explained by existing knowledge of the genetic causes of the WS phenotype<sup>18,24–26</sup>. **c–h**, Colonies of (c) original ancestral genotype, (d) immediate ancestor of bet-hedging genotype, (e) bet-hedging genotype, (f) original ancestral genotype with mutated *carB*, (g) immediate ancestor of bet-hedging genotype with mutated *carB* and (h) bet-hedging genotype with wild-type *carB*. Colonies were photographed at the same magnification.

and competed the resulting genotype against  $1A^4$ . This showed that the *carB* mutation increases the fitness of  $1A^4$  (one-sample *t*-test,  $n = 9$ ,  $P = 0.0013$ ; Fig. 4). Moreover,  $1B^4$  with a wild-type *carB* allele had a lower fitness relative to  $1B^4$ , confirming that the *carB* mutation was essential for the increased fitness of this genotype (one-sample *t*-test,  $n = 9$ ,  $P = 0.0004$ ; Fig. 4). Together, these results indicate that the fitness effect of the *carB* mutation is sufficient to explain the emergence of  $1B^4$  to a detectable frequency.

Colony switching was caused by a single point mutation but nonetheless took nine rounds of selection to evolve. This led us to question the importance of the previously fixed mutations<sup>28</sup>. To study this, one round of the evolutionary experiment was repeated from both  $1A^4$  and the original ancestor ( $1A^0$ ). Colony switching evolved from  $1A^4$  (3 from 36 replicates), but never from the ancestral genotype (0 from 138 replicates), indicating that these genotypes differed in their capacity to give rise to colony switching (two-tailed Fisher's exact test,  $P = 0.0083$ ).

The reliance on previously fixed mutations might stem from epistatic interactions essential for the *carB* mutation to cause colony switching or to confer the requisite fitness benefit in static microcosms.



**Figure 4 | Relative fitness of evolved and engineered strains.** Box plot of the fitness of a bet-hedging genotype relative to its immediate ancestor ( $1B^4$  versus  $1A^4$ ), and the fitness effects of mutated *carB* in the immediate ancestor of the bet-hedging genotype ( $1A^4$  mut versus  $1A^4$ ), wild-type *carB* in  $1B^4$  ( $1B^4$  WT versus  $1B^4$ ) and mutated *carB* in the original ancestor ( $1A^0$  mut versus  $1A^0$ ). All fitness assays were performed in static microcosms. Values greater than 1 (dashed line) indicate a higher relative fitness of the first competitor. Key: median (horizontal lines in boxes), interquartile range (boxes), 90th and 10th percentiles (vertical bars), significant deviation from 1 (\*\*\*  $P \leq 0.001$ , \*\*  $P \leq 0.01$ ; see text for statistics).

We distinguished between these hypotheses by introducing the *carB* mutation in  $1A^0$ . In this background it did cause colony switching (Fig. 3f) but appeared not to confer a significant fitness increase (Fig. 4). The latter was confirmed by a direct statistical comparison with the effect of the *carB* mutation in  $1A^4$  (analysis of covariance, the larger variance of the  $1A^0$  data is explained by a covariate), which identified a significant epistatic interaction (analysis of deviance,  $F(1,16) = 8.536$ ,  $P = 0.01$ ), and indicated that the *carB* mutation is beneficial in  $1A^4$  but deleterious in  $1A^0$  (Supplementary Note 4). We conclude that the evolutionary history of  $1A^4$  'set the stage' for the evolution of stochastic colony morphology switching by altering the relative fitness effect of the *carB* mutation.

Owing to the historical nature of the evolutionary process, the origins of most adaptive phenotypes are obscure<sup>11</sup>. Here we have provided a mechanistic account of the adaptive evolution of a widespread trait<sup>4–10</sup>. Bet hedging arose as an adaptation to fluctuating selection imposed by an exclusion rule and bottleneck, two population processes that are likely to play a key role in the evolution of stochastic phenotype switching in nature. Insight into the underlying molecular details reveals how evolution tinkered with central metabolism to generate a strategy that could reasonably—one might think—have taken tens of thousands of generations to evolve. The rapid and repeatable evolution of bet hedging during our experiment suggests it may have been among the earliest evolutionary solutions to life in variable environments, perhaps even preceding the evolution of environmentally responsive mechanisms of gene regulation.

## METHODS SUMMARY

*P. fluorescens* SBW25 (ref. 23) was grown in glass microcosms containing liquid medium. During each selection round, populations were propagated by transfer of a sample to a fresh microcosm. Parallel with each transfer, populations were checked for the presence of cells that formed new colony types. Identification of a colony different from that with which the selection round had been started marked the end of a selection round. Cells from these colonies were stored, and used to found the next selection round in the opposing environment. This procedure was repeated 15 times. Control lines were continually selected in shaken microcosms and always propagated by transfer of cells from colonies of the ancestral type. Relative fitness was measured in competition assays, using colony morphology or a neutral marker<sup>29</sup> to distinguish genotypes, and expressed as the ratio of Malthusian parameters<sup>30</sup>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 7 July; accepted 15 September 2009.

- Cohen, D. Optimizing reproduction in a randomly varying environment. *J. Theor. Biol.* 12, 119–129 (1966).

2. Seger, J. & Brockmann, H. J. in *Oxford Surveys in Evolutionary Biology*, Vol. 4 (eds Harvey, P. & Partridge, L.) 182–211 (Oxford Univ. Press, 1987).
3. Slatkin, M. Hedging one's evolutionary bets. *Nature* **250**, 704–705 (1974).
4. Ackermann, M. *et al.* Self-destructive cooperation mediated by phenotypic noise. *Nature* **454**, 987–990 (2008).
5. Danforth, B. N. Emergence dynamics and bet hedging in a desert bee, *Perdita portalis*. *Proc. Natl Acad. Sci. USA* **266**, 1985–1994 (1999).
6. Friedenberg, N. A. Experimental evolution of dispersal in spatiotemporally variable microcosms. *Ecol. Lett.* **6**, 953–959 (2003).
7. Hairston, N. G. & Munns, W. R. The timing of copepod diapause as an evolutionary stable strategy. *Am. Nat.* **123**, 733–751 (1984).
8. Moxon, E. R., Rainey, P. B., Nowak, M. A. & Lenski, R. E. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.* **4**, 24–33 (1994).
9. Tonegawa, S. Somatic generation of antibody diversity. *Nature* **302**, 575–581 (1983).
10. Venable, D. L. Bet hedging in a guild of desert annuals. *Ecology* **88**, 1086–1090 (2007).
11. Gould, S. J. & Lewontin, R. C. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. B* **205**, 581–598 (1979).
12. Kussell, E. & Leibler, S. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**, 2075–2078 (2005).
13. Wolf, D. M., Vazirani, V. V. & Arkin, A. P. Diversity in times of adversity: probabilistic strategies in microbial survival games. *J. Theor. Biol.* **234**, 227–253 (2005).
14. Meyers, L. A. & Bull, J. J. Fighting change with change: adaptive variation in an uncertain world. *Trends Ecol. Evol.* **17**, 551–557 (2002).
15. Acar, M., Mettetal, J. T. & van Oudenaarden, A. Stochastic switching as a survival strategy in fluctuating environments. *Nature Genet.* **40**, 471–475 (2008).
16. Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. Bacterial persistence as a phenotypic switch. *Science* **305**, 1622–1625 (2004).
17. Rainey, P. B. & Travisano, M. Adaptive radiation in a heterogeneous environment. *Nature* **394**, 69–72 (1998).
18. Spiers, A. J., Kahn, S. G., Bohannon, J., Travisano, M. & Rainey, P. B. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. *Genetics* **161**, 33–46 (2002).
19. Giddens, S. R. *et al.* Mutational activation of niche-specific genes provides insight into regulatory networks and bacterial function in a complex environment. *Proc. Natl Acad. Sci. USA* **104**, 18247–18252 (2007).
20. Stevenson, G., Andrianopoulos, K., Hobbs, M. & Reeves, P. R. Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *J. Bacteriol.* **178**, 4885–4893 (1996).
21. Albert, T. J. *et al.* Mutation discovery in bacterial genomes: metronidazole resistance in *Helicobacter pylori*. *Nature Methods* **2**, 951–953 (2005).
22. Bentley, D. R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59 (2008).
23. Silby, M. W. *et al.* Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. *Genome Biol.* **10**, R51 (2009).
24. Bantinaki, E. *et al.* Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. III. Mutational origins of wrinkly spreader diversity. *Genetics* **176**, 441–453 (2007).
25. Goymer, P. *et al.* Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. II. Role of the GGDEF regulator WspR in evolution and development of the wrinkly spreader phenotype. *Genetics* **173**, 515–526 (2006).
26. McDonald, M. J., Gehrig, S. M., Meintjes, P. L., Zhang, X. X. & Rainey, P. B. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. IV. Genetic constraints guide evolutionary trajectories in a parallel adaptive radiation. *Genetics* (in the press).
27. Mergeay, M., Gigot, D., Beckman, J., Glansdorff, N. & Pierard, A. Physiology and genetics of carbamoylphosphate synthesis in *Escherichia coli* K-12. *Mol. Gen. Genet.* **133**, 299–316 (1974).
28. Blount, Z. D., Borland, C. Z. & Lenski, R. E. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **105**, 7899–7906 (2008).
29. Zhang, X. X. & Rainey, P. B. Construction and validation of a neutrally-marked strain of *Pseudomonas fluorescens* SBW25. *J. Microbiol. Methods* **71**, 78–81 (2007).
30. Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am. Nat.* **138**, 1315–1341 (1991).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank T. F. Cooper, M. R. Goddard and D. Refardt for discussion; M. Gray, A. Hurman and G. E. M. Lamers for technical assistance; M. Ackermann, P. M. Brakefield, T. Fukami, S. Rossell and B. J. Zwaan for comments on the manuscript; T. J. M. Van Dooren for statistical advice; F. Bertels for computational analysis of Solexa data; and E. Libby for theoretical insight. This work was supported by the Marsden Fund Council from government funding administered by the Royal Society of New Zealand. H.J.E.B. is supported by a Veni Fellowship from The Netherlands Organisation for Scientific Research (NWO). J.G. was supported by a Bright Futures Scholarship from the New Zealand Foundation for Research, Science and Technology. C.K. was supported by a Feodor Lynen Fellowship from the Alexander von Humboldt Foundation, Germany. G.C.F. is supported by a Postdoctoral Fellowship from the New Zealand Foundation for Research, Science and Technology.

**Author Contributions** H.J.E.B. and P.B.R. conceived the research and wrote the manuscript. H.J.E.B. conducted the main selection experiment, examined the relation between cell and colony morphology, investigated reversible switching, performed genome analysis and contributed to allelic replacements. J.G. determined Cap<sup>+</sup> cell proportions, performed transposon mutagenesis and integration-site identification, confirmed and ordered the mutations, performed allelic replacements, and contributed to fitness assays. C.K. and G.C.F. performed fitness assays. All authors commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to H.J.E.B. ([h.j.e.beaumont@biology.leidenuniv.nl](mailto:h.j.e.beaumont@biology.leidenuniv.nl)).



## METHODS

**Strain and medium.** *P. fluorescens* SBW25 (ref. 23) was grown at 28 °C in 25-ml glass microcosms containing 6 ml King's Medium B with loose caps, and on King's Medium B–agar plates for 48 h.

**Selection regime.** Populations were founded with approximately  $10^7$  cells from –80 °C glycerol stocks prepared from cells derived from a single colony. For the first round, populations were grown in static microcosms and propagated by transfer of a mixed sample (6 µl) to a fresh microcosm at 72-h intervals of approximately ten generations. Parallel with each transfer, samples were spread onto plates to screen for colonies with a different morphology (approximately 1,000 screened, 500 minimum). Fuzzy spreader genotypes<sup>17</sup> did not respond to selection (H.J.E.B., unpublished observations) and were omitted. Upon detection of new colony types, a single colony of the numerically dominant new type was streaked to single colonies on a control plate to confirm heritability of colony morphology, and to ensure a single-cell bottleneck. Emergence of a colony with a different morphology

marked the end of a round. Cells of these colonies were taken from control plates, grown in a shaken microcosm (16 h) and stored at –80 °C. Fifteen additional selection rounds (alternately shaken at 150 r.p.m. and static) were performed, each founded with the cells from the colony selected in the previous round. Control lines were continually selected in shaken microcosms (for the same number of rounds as the cognate selection lines) and maintained under stabilizing selection for the ancestral colony type by continual selection of a numerically dominant smooth<sup>17</sup> colony.

**Fitness assays.** Cells taken from single colonies (grown from –80 °C stocks) were pre-conditioned by cultivation in shaken microcosms for 24 h. Competitions were inoculated with approximately  $5 \times 10^6$  cells of each competitor, and incubated for 72 h. Competitor frequencies were determined by plating at 0 h and 72 h. Genotypes were distinguished by colony morphology or a neutral marker<sup>29</sup>. Relative fitness was expressed as the ratio of Malthusian parameters<sup>30</sup>.

## LETTERS

# Effector T cell interactions with meningeal vascular structures in nascent autoimmune CNS lesions

Ingo Bartholomäus<sup>1\*</sup>, Naoto Kawakami<sup>1\*</sup>, Francesca Odoardi<sup>1,2,3</sup>, Christian Schläger<sup>1,2</sup>, Djordje Miljkovic<sup>1</sup>, Joachim W. Ellwart<sup>4</sup>, Wolfgang E. F. Klinkert<sup>1</sup>, Cassandra Flügel-Koch<sup>5</sup>, Thomas B. Issekutz<sup>6</sup>, Hartmut Wekerle<sup>1</sup> & Alexander Flügel<sup>1,2,3</sup>

The tissues of the central nervous system are effectively shielded from the blood circulation by specialized vessels that are impermeable not only to cells, but also to most macromolecules circulating in the blood. Despite this seemingly absolute seclusion, central nervous system tissues are subject to immune surveillance and are vulnerable to autoimmune attacks<sup>1</sup>. Using intravital two-photon imaging in a Lewis rat model of experimental autoimmune encephalomyelitis, here we present in real-time the interactive processes between effector T cells and cerebral structures from their first arrival to manifest autoimmune disease. We observed that incoming effector T cells successively scanned three planes. The T cells got arrested to leptomeningeal vessels and immediately monitored the luminal surface, crawling preferentially against the blood flow. After diapedesis, the cells continued their scan on the abluminal vascular surface and the underlying leptomeningeal (pial) membrane. There, the T cells encountered phagocytes that effectively present antigens, foreign as well as myelin proteins. These contacts stimulated the effector T cells to produce pro-inflammatory mediators, and provided a trigger to tissue invasion and the formation of inflammatory infiltrations.

Autoimmune inflammation of the central nervous system (CNS) was induced by intravenous transfer of CD4<sup>+</sup> effector T cells reactive against the myelin component, myelin basic protein (MBP)<sup>2</sup>. The encephalitogenic effector T cells had been retrovirally transduced to express green fluorescent protein (T<sub>MBP-GFP</sub> cells) as a marker for identification *in situ* and isolation *ex vivo*<sup>3</sup>. We pursued the fate of these cells within their target organ throughout the 3–4 days between their transfer and disease onset (Supplementary Fig. 1a) by scanning by intravital two-photon microscopy the lower spinal cord tissues from the leptomeninges to the adjacent lumbar fasciculus at a depth of ~200 µm (Supplementary Fig. 1b)<sup>4</sup>.

The first T<sub>MBP-GFP</sub> cells appeared in the CNS before onset of clinical experimental autoimmune encephalomyelitis (EAE), on days 1–2.5 after transfer. In line with earlier work<sup>5,6</sup>, the arriving cells were found to be restricted to subarachnoidal areas (Supplementary Fig. 2a, b). Notably, the cells remained strictly in close association with pial blood vessels, crawling on surfaces within the outline of the vessels (Fig. 1a and Supplementary Movie 1). This behaviour was specific to the CNS. In contrast, in peripheral organs, for example in peripheral nerves, muscle or subcutaneous tissue, T<sub>MBP-GFP</sub> cells mainly rolled along the inner surface of the vessels (Fig. 1b and Supplementary Movie 2). Several observations indicated that the freshly arrived vessel-associated T cells crawled on the inner vascular surface: first, we often saw T cells getting captured to the vessel wall before migrating along the vascular

surface. Conversely, other vessel-associated, crawling T<sub>MBP-GFP</sub> cells detached and were washed away by the blood stream (Supplementary Fig. 3a, b and Supplementary Movie 3). Second, in single-fluorescence images of fluorescent dextran-filled vessels, intraluminal T cells displaced the fluorescent dextran-fill, creating dark shadows within the vascular volume (Supplementary Fig. 3c)<sup>7</sup>. Third, and most directly, we traced T<sub>MBP-GFP</sub> cells passing through the vessel walls (Fig. 1c and Supplementary Movie 4).

Crawling on vascular endothelia was recently recognized as stereotypic behaviour of polymorphonuclear leukocytes<sup>8</sup>, monocytes<sup>7</sup> and natural killer T cells<sup>9</sup>, but has not previously been seen with (auto-immune) T cells *in situ*<sup>7</sup>. Intraluminal crawling seems to be characteristic of post-activated migration-competent effector T cells directly before and during the build-up of clinical and histological EAE<sup>10</sup>. It contrasts with freshly *in vitro*-activated encephalitogenic T cells injected into naive hosts, which penetrate CNS vessels after abrupt arrest without rolling or crawling<sup>11</sup>.

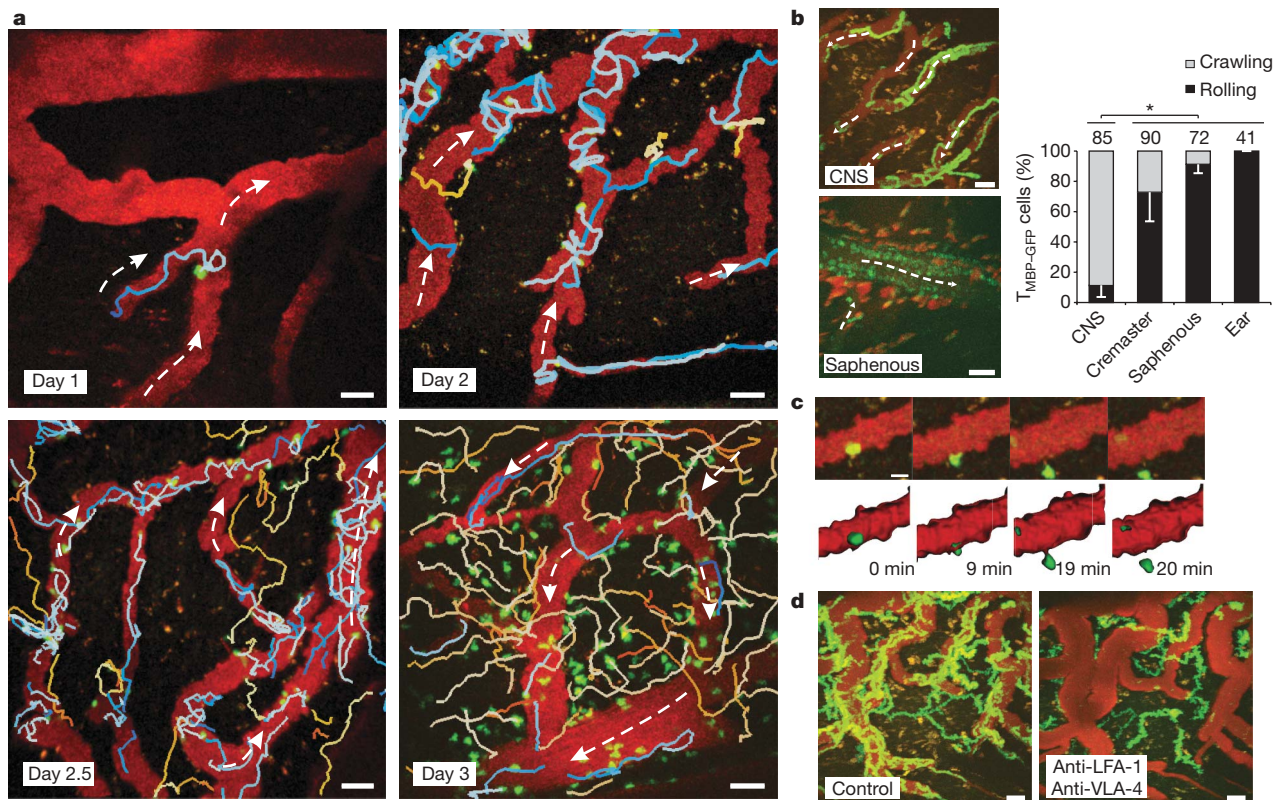
Intraluminal crawling of encephalitogenic T cells is not a peculiarity of rat EAE, nor is it an artefactual feature of EAE transferred by cultured T cells: mouse effector T cells similarly crawled within CNS vessels during initial stages of EAE (Supplementary Fig. 4a). Moreover, T-cell crawling was also seen in actively induced EAE, with autoaggressive memory T cells, which had been embedded in the rat immune system over months<sup>12</sup> and were activated *in vivo* by immunization against MBP (Supplementary Fig. 4b).

Cell locomotion on the vascular surface probably involves the formation of reversible, punctate contacts with substrates by complementary cell adhesion molecules<sup>13</sup>. In fact, the integrin LFA-1 was implicated in the crawling of monocytes<sup>7</sup>, whereas VLA-4 acts as a key adhesion molecule in T-cell passage through the blood–brain barrier in EAE<sup>14</sup>. Both integrins had a role in T<sub>MBP-GFP</sub> cell locomotion, although with different weight. Monoclonal antibodies against VLA-4 stopped crawling and transmigration of the T<sub>MBP-GFP</sub> cells into the CNS parenchyma, a process that took a few hours (Supplementary Fig. 5a, b), whereas anti-LFA-1 antibody treatment per se had no clear effect. In combination, however, anti-LFA-1 and anti-VLA-4 antibodies detached crawling T cells from the luminal walls almost instantaneously (Fig. 1d and Supplementary Movie 5). As expected, anti-VLA-4 treatment strongly reduced the development of neurological deficits (Supplementary Fig. 5c). In contrast, anti-LFA-1 was not protective, nor did it add to the therapeutic effect of anti-VLA-4.

CNS invasion by T<sub>MBP-GFP</sub> cells escalated from day 2 after transfer. The numbers of intraluminal T<sub>MBP-GFP</sub> cells increased steeply, and numerous T cells passed to the outer side of the pial vessels (Figs 1a,

<sup>1</sup>Max Planck Institute for Neurobiology, 82152 Martinsried, Germany. <sup>2</sup>Department of Neuroimmunology, Institute for Multiple Sclerosis Research, Gemeinnützige Hertie-Stiftung and University Medical Centre Göttingen, 37073 Göttingen, Germany. <sup>3</sup>Institute for Immunology, Ludwig-Maximilians-University, 80336 Munich, Germany. <sup>4</sup>Institute for Experimental Hematology, Helmholtz Centre, 81377 Munich, Germany. <sup>5</sup>Institute for Anatomy 2, Friedrich-Alexander-University, 91054 Erlangen, Germany. <sup>6</sup>Division of Immunology, Department of Pediatrics, Dalhousie University, Halifax, Nova Scotia B3K 6R8, Canada.

\*These authors contributed equally to this work.



**Figure 1 | Intravital two-photon recording of T<sub>MBP</sub>-GFP cells entering the leptomeninges.** **a**, Migratory paths (30 min) of intraluminal (blue lines) or extravasated (yellow lines) T<sub>MBP</sub>-GFP cells (green) in the lumbar spinal cord of rats at the indicated time points are shown. Colour intensities increase with time. Arrows indicate the blood flow. Blood vessels are labelled in red. Representative data of at least six experiments (obtained from individual animals) per time point are shown. Scale bars, 50  $\mu$ m. **b**, T<sub>MBP</sub>-GFP cells (green) within vessels of the peripheral nervous system or the CNS 2.5 days after transfer. Time projections of 30 min are shown. Rolling cells appear as several dots (Methods). Representative pictures are from three (saphenous nerve) and six (CNS) independent experiments, respectively. Scale bars, 50  $\mu$ m. Bar diagram: percentage of intraluminally crawling or rolling

T<sub>MBP</sub>-GFP cells in vascular beds of the leptomeninges (CNS), cremaster muscle, saphenous nerve or in subcutaneous ear tissue 3 days after transfer. Results represent the mean and s.d. of at least three independent experiments per organ. Numbers at top indicate the numbers of analysed cells.  $P < 0.05$  (Mann–Whitney U test). **c**, The diapedesis of a T<sub>MBP</sub>-GFP cell (green) through a leptomeningeal vessel (red) at day 2 after transfer is depicted. Scale bar, 10  $\mu$ m. **d**, Detachment of intraluminally crawling T<sub>MBP</sub>-GFP cells after co-application of anti-LFA-1 and anti-VLA-4 antibodies 48 h after transfer. Migratory tracks (green, 30 min duration) of T<sub>MBP</sub>-GFP cells before (control) or 5 min after intravenous antibodies infusion are shown. A representative experiment of more than three independent experiments is shown. Scale bars, 50  $\mu$ m.

2a, Supplementary Fig. 2c and Supplementary Movie 6). Most of the extravasated T cells remained initially within the vascular environment (Figs 1a, 2a, Supplementary Figs 2c, 6a and Supplementary Movie 7) before detaching and crawling on the surface of the neuropil seemingly at random (Fig. 1a, Supplementary Fig. 2d, e and Supplementary Movie 7).

Thus, the infiltrating T<sub>MBP</sub>-GFP cells crawled successively on three distinct cellular planes, but with slightly distinct velocities. Whereas the average velocity of intraluminal T cells was  $12.5 \pm 3.4 \mu\text{m min}^{-1}$ , with or against blood flow, extravascular T cells were slightly slower ( $11.3 \pm 3.3 \mu\text{m min}^{-1}$ ; Fig. 2b and Supplementary Fig. 3d). The T cells principally migrated in all directions, although, notably, intraluminal T cells moved preferentially against the blood flow (42% versus 19%; Fig. 2c). On average, T cells spent  $\sim 15$  min crawling, although a substantial fraction (5–10%) continued their hike in the vascular bed over the entire observation time of more than 30 min (Supplementary Fig. 3e).

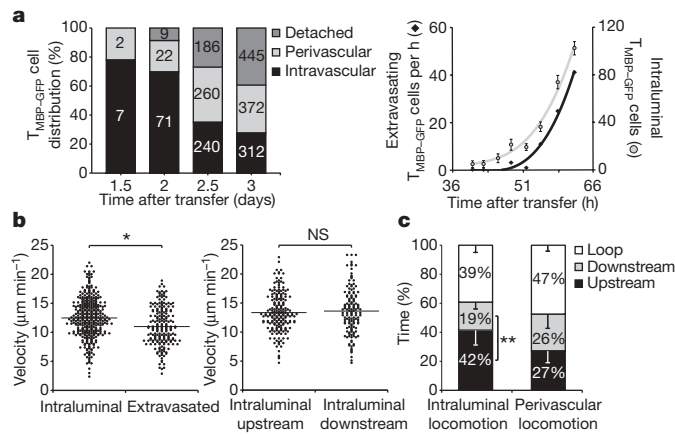
Diapedesis and scanning of the leptomeninges by the T<sub>MBP</sub>-GFP cells were the prelude to actual invasion of the CNS tissue. The T cells intruded into the Virchow–Robin spaces around the radial vessels that dive into the spinal cord parenchyma (Supplementary Fig. 6b). Submeningeal and perivascular effector T<sub>MBP</sub>-GFP cells invaded the adjacent CNS parenchyma—a process that coincided with the onset of clinical symptoms on day 3 after transfer (Supplementary Figs 1a and 2d). One day later, at the peak of disease symptoms, the effector T cells were found dispersed throughout the white and grey spinal

cord matter (Supplementary Fig. 2e). This immigration process was confirmed by quantitative cytofluorometric analyses (Supplementary Fig. 6c).

So far, we have considered the behaviour of CNS invading encephalitogenic T<sub>MBP</sub>-GFP cells. Another consideration is the action of T cells reactive against non-brain antigens, for example, ovalbumin-reactive T cells (T<sub>OVA</sub>-GFP cells). Similar to T<sub>MBP</sub> cells, intravenously transferred T<sub>OVA</sub>-GFP cells circulated in the blood and rolled along peripheral vessel walls. Notably, T<sub>OVA</sub>-GFP cells also accumulated within the leptomeninges. Similar to T<sub>MBP</sub> cells in their early state of arrival, T<sub>OVA</sub>-GFP cells crawled within pial vessels, also preferentially against the blood flow, at intraluminal and extravascular velocities similar to those of T<sub>MBP</sub> cells (Supplementary Fig. 7a–c). In contrast to T<sub>MBP</sub> cells, T<sub>OVA</sub>-GFP cells spent less time crawling within vessels (Supplementary Fig. 7d), and their numbers in the leptomeninges were lower. Most importantly, T<sub>OVA</sub>-GFP cells did not enter the CNS parenchyma in significant numbers (Supplementary Fig. 11a).

Up to now, we showed that T<sub>MBP</sub>-GFP cells scan the CNS vessel-associated milieu on at least three levels, but the objects of the apparent search remained unknown. Intraluminal T cells, like human monocytes<sup>15</sup>, might screen the vascular endothelium to spot sites favourable to intercellular or transcellular diapedesis<sup>16</sup>. However, once extravasated, T<sub>MBP</sub>-GFP cells face other signals. We noted that T cells transgressing the pial vascular wall made contact with perivascular/meningeal phagocytes, which were identified either by intrathecally infused Texas-red-tagged dextran<sup>17</sup> or, in bone-marrow chimaeric



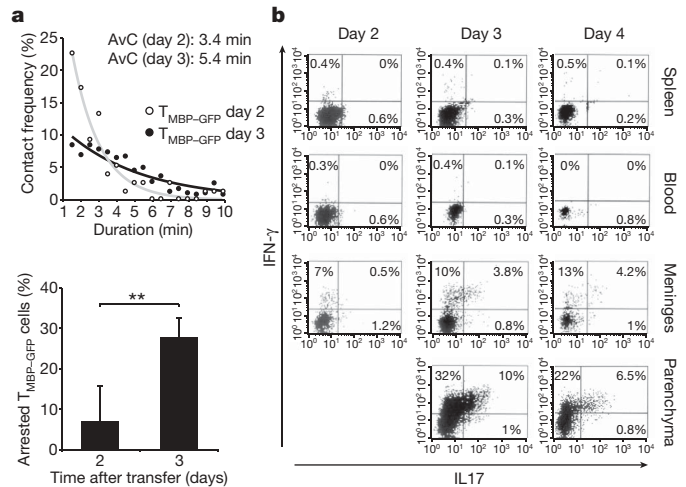


**Figure 2 | T<sub>MBP-GFP</sub> cell locomotion within meningeal CNS structures.**

**a**, T<sub>MBP-GFP</sub> cell locomotion within the lumen (intravascular) or the outer sides (perivascular) of leptomeningeal vessels, or of T cells migrating independently from vessels (detached) was quantified at the indicated time points (left). Total numbers of T<sub>MBP-GFP</sub> cells within an area of 1.2 mm<sup>2</sup> are indicated. Data were obtained by intermittent recording over an observation period of 35 h. Kinetics of T-cell crawling and diapedesis are shown on the right. Intraluminal crawling (circles) and extravasating T<sub>MBP-GFP</sub> cells were quantified at the indicated time points. Data represent mean and s.d. The diapedesis rate (diamonds) was calculated from the increase of extravasated T<sub>MBP-GFP</sub> cells per hour. Mean numbers per imaging field (0.36 mm<sup>2</sup>) are shown. A representative experiment of three independent experiments is shown. **b**, Locomotion characteristics of effector T cells. Average velocities of T<sub>MBP-GFP</sub> cells moving within or on the outside of leptomeningeal vessels (days 2–3 after transfer) are shown. Data include 365 and 201 cells (10,299 and 7,455 displacements) from more than five independent experiments for intraluminal and extravasated T<sub>MBP-GFP</sub> cells, respectively (left). Average velocities of intraluminal crawling T<sub>MBP-GFP</sub> cells moving with or against the blood flow ( $n = 140$  and  $193$  cells; right). Dots represent average velocities of individually tracked cells. \* $P < 0.05$ ; NS, not significant (two-tailed unpaired  $t$ -test). **c**, Orientation of intraluminal or perivascularly crawling T<sub>MBP-GFP</sub> cells. Loop, T cells moving at an angle of  $\geq 45^\circ$  of the vessel axis; downstream or upstream, T cells crawling along the axis of the vessels with or against the blood flow, respectively. Data include 218 intraluminal and 56 extravasated T<sub>MBP-GFP</sub> cells (6,570 and 1,605 displacements, respectively) from at least four independent experiments recorded at 2–2.5 days after transfer. Means and s.d. are shown. \*\* $P < 0.01$  ( $\chi^2$ ). Standard deviation refers to mean values of individual experiments.

rats, by transgenic GFP<sup>18</sup>. These cells were located strategically around vessels and within the pial membrane, monitoring the environment with their cellular processes (Supplementary Fig. 8a and Supplementary Movie 8). The phagocytes shared some features with dendritic cells, and others with macrophages<sup>19,20</sup>. They were functional antigen-presenting cells, expressing major histocompatibility complex (MHC) class II determinants (60%), and efficiently presenting not only exogenous proteins, but also endogenous CNS autoantigen to T cells *in vitro* (Supplementary Fig. 8b–d)<sup>21</sup>.

Having crossed the vascular walls, T<sub>MBP-GFP</sub> cells formed diverse contacts: some were short lasting (<10 min), often serially involving several phagocytes, whereas others lasted for longer periods of time (Fig. 3a, Supplementary Fig. 9a and Supplementary Movie 9). Notably, the duration of these contacts and the numbers of long-lasting arrests increased with ongoing inflammation (Fig. 3a). The extravascular T<sub>MBP-GFP</sub> cells became strongly reactivated, with upregulated pro-inflammatory cytokines (interferon (IFN)- $\gamma$ , interleukin (IL)17, tumour necrosis factor (TNF)- $\alpha$ , IL2), proteases (matrix metalloproteinase (MMP)9, 14), chemokines/chemokine receptors (CCL5, CCR5, CXCR3, CXCR4) and surface activation markers (OX40 antigen, IL2 receptor). The mitotic activity, however, remained low as indicated by the low expression of cyclins B1, B2, D3, E and the lack of BrdU uptake. This gene expression profile was essentially maintained through infiltration into the CNS parenchyma (Fig. 3b and Supplementary Fig. 9b–e).

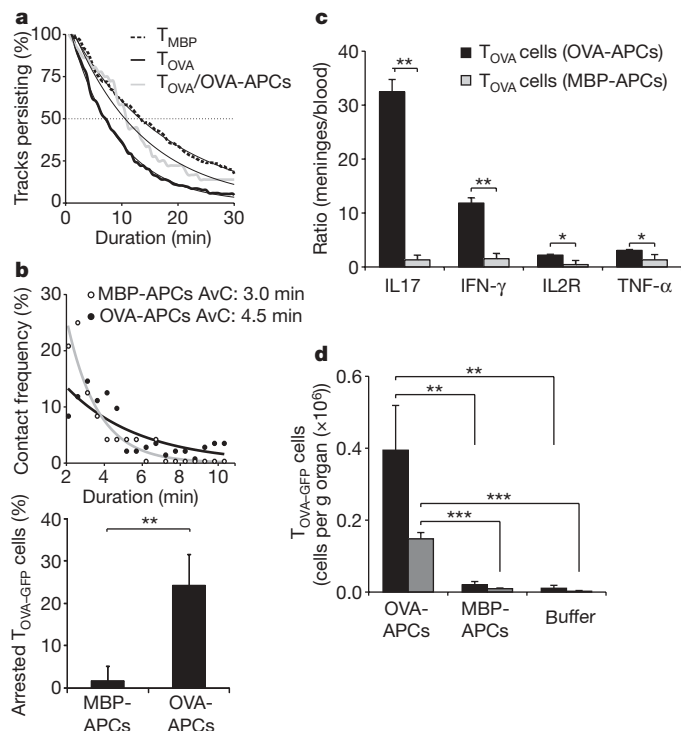


**Figure 3 | Effector T cells leaving the leptomeningeal vessels encounter perivascular/meningeal antigen-presenting cells.**

**a**, Quantification of T cell–antigen-presenting cell contacts. Serial short-lasting (<10 min) contacts of T<sub>MBP-GFP</sub> cells at days 2 and 3 after transfer (top). Frequencies of individual contact durations are shown. AvC, average duration of T<sub>MBP-GFP</sub>–antigen-presenting cell contacts. There were 74 and 454 contacts analysed at days 2 and 3, respectively. Representative data of at least three experiments per time point are shown. Numbers of stable arrests (>10 min) at days 2 and 3 after transfer are shown in the bottom panel. Data include 80 cells from seven experiments (day 2) and 375 cells from six experiments (day 3). Means and s.d. are shown. \*\* $P < 0.01$  (Mann–Whitney U test). **b**, Pro-inflammatory cytokine expression of T<sub>MBP-GFP</sub> cells within CNS structures. IFN- $\gamma$  and IL17 protein expression of T<sub>MBP-GFP</sub> cells at the indicated time points were measured by cytofluorometry. The numbers indicate the percentage of cells in each quadrant. Representative dot plots of at least four independent experiments are shown.

Encounters with antigen-presenting phagocytes activate the T<sub>MBP-GFP</sub> cells, but questions remain as to the consequences of these interactions, in particular, whether they are essential prerequisites for the invasion of the compact CNS tissue. A study of ovalbumin as ‘neoautoantigen’ seems to support this hypothesis. In the absence of antigen, the few brain antigen-ignorant T<sub>OVA-GFP</sub> cells that attached to the luminal vascular surface, and crossed the vessel wall, formed transient contacts with perivascular/meningeal phagocytes (Fig. 4b), but they did not upregulate activation markers nor pro-inflammatory cytokines (Fig. 4c). However, contact with OVA-presenting phagocytes changed the behaviour of T<sub>OVA-GFP</sub> cells profoundly. T<sub>OVA-GFP</sub> cells that had crossed the vessel walls formed long-lasting contacts with OVA-pulsed phagocytes followed by upregulating of membrane activation markers and pro-inflammatory cytokines (Fig. 4b, c and Supplementary Fig. 10a). Activation of T<sub>OVA-GFP</sub> cells had drastic consequences. Unexpectedly, it was followed by a steep increase in intraluminal crawling T<sub>OVA-GFP</sub> cells and by a considerable prolongation of crawling tracks (Fig. 4a and Supplementary Fig. 10b). Furthermore, some of the activated T<sub>OVA-GFP</sub> cells intruded deeply into the CNS parenchyma, which went along with the development of mild clinical signs such as transient weight loss (Fig. 4d and Supplementary Fig. 10c). A similar parenchymal distribution pattern of T<sub>OVA-GFP</sub> cells occurred after co-transfer with T<sub>MBP</sub> cells—that is, when T<sub>OVA-GFP</sub> cells met an activated CNS milieu (Supplementary Fig. 11a, b). It seems that local extravascular inflammation caused by the re-activation of T<sub>OVA-GFP</sub> cells influences the intraluminal locomotion of effector T cells, probably by retrograde activation of the CNS vascular endothelia, and at the same time triggers their invasion into the CNS parenchyma.

Our intravital two-photon study sheds new light on two important aspects of the immune reaction: immune surveillance of the CNS, and in a broader context, tissue-specific autoimmune disease. We describe that autoimmune lesions are initiated around pial veins, with the incoming T cells systematically scanning first the inner, then



**Figure 4 | Brain-antigen-ignorant effector T cells invade the CNS tissue after antigen-driven activation.** **a**, Intensified intraluminal crawling of  $T_{OVA-GFP}$  cells after contact with OVA-pulsed meningeal antigen-presenting cells (OVA-APCs). Fractions of  $T_{MBP-GFP}$  cells,  $T_{OVA-GFP}$  cells and of  $T_{OVA-GFP}$  cells 3 h after intrathecal transfer of OVA-APCs are plotted as a function of their crawling durations. There were 180, 221 and 36 individual tracks analysed, respectively. Data were obtained from seven intravital experiments. Trend lines are shown with thin black lines. **b**, Serial short-lasting (<10 min) contacts of  $T_{OVA-GFP}$  cells 14 h after OVA-APC or MBP-APC transfer (top). Frequencies of individual contact durations are shown. AvC, average duration of  $T_{OVA-GFP}$ -APC contacts. There were 221 and 25 contacts from three intravital recordings/treatment analysed, respectively. Numbers of stable arrests (>10 min) after OVA-APC or MBP-APC transfer (bottom). A representative data set of at least three experiments is shown. Data include 26 and 387 cells for MBP- and OVA-pulsed APCs, respectively. Values indicate means and s.d.  $**P < 0.01$  (Mann–Whitney U test). **c**, Regulation of pro-inflammatory cytokines in the presence of OVA- or MBP-APCs (quantitative PCR of *ex vivo* purified  $T_{OVA-GFP}$  cells). Values represent means and s.d.  $*P < 0.05$ ,  $**P < 0.01$  (two-tailed unpaired *t*-test). **d**, Cytofluorometric quantification of  $T_{OVA-GFP}$  cells within meninges (black bars) or CNS parenchyma (grey bars) 24 h after APC transfer. Representative data of seven independent experiments are shown.  $**P < 0.01$ ,  $***P < 0.001$  (one-way analysis of variance (ANOVA) (Holm–Sidak method)).

the outer vascular surfaces and the subarachnoid space on at least three distinct levels (Supplementary Fig. 12). These observations establish the leptomeningeal, transvascular route as the main portal used by encephalitogenic T cells in early EAE, although contributions by alternative pathways, for example, by the choroid plexus, are not excluded<sup>22</sup>. Activation of the brain autoreactive effector T cells by contacts with antigen-presenting perivascular/meningeal phagocytes sets the stage for their entry into the underlying CNS tissue and the recruitment of further immune cells<sup>23–26</sup>. These observations detail initial processes of immune surveillance, and hence are relevant for understanding CNS infection and tumour growth. We expect that the structures involved in T-cell interactions with endothelium, perivascular phagocytes and matrix components qualify as targets for new and selective therapies in brain autoimmune disease.

## METHODS SUMMARY

**EAE model, T cells.** Lewis rat (from the breeding facilities of the Max Planck Institutes) adoptive transfer EAE was induced by intravenous transfer of  $5 \times 10^6$

MBP-specific  $CD4^+$  T cells<sup>2</sup>. GFP-transgenic (GFPtg) rats (GFP expression driven by the ubiquitin C promoter) were provided by D. Baltimore<sup>27</sup>. All experiments were conducted according to Bavarian and Lower Saxonian State regulations for animal experimentation and approved by the appropriate authorities. MBP-specific and OVA-specific  $CD4^+$  T cells ( $T_{MBP}$  and  $T_{OVA}$ ) were retrovirally transduced to express the fluorescent protein genes<sup>3</sup>.

**Live imaging.** Anaesthesia of animals and preparation of spinal windows were described previously<sup>4</sup>. For time-lapse two-photon laser-scanning microscopy a Leica SP2 microscope (Leica Microsystems) equipped with a  $\times 20$  water immersion objective (Olympus, numerical aperture (NA) 0.95) was combined with a 10W-Millenia/Tsunami femto-second laser (Spectra Physics)<sup>28</sup>. Images were generally obtained in 32-s time intervals for the indicated time periods. Movies are shown with a speed of five frames per second. Fluorescence labelling of vessels was achieved by intravenous infusion of Texas-red- or FITC-conjugated dextran (70 kDa). Perivascular/meningeal macrophages were marked by intrathecal injection of fluorescent dextran conjugates.

**Flow cytometry, FACS, immunofluorescence, proliferation assays and quantitative PCR reactions.** Cells were isolated *ex vivo* or retransferred intrathecally using established techniques<sup>4,29</sup>. Functional tests to evaluate the transcriptome and cytokine profile of T cells were performed as described previously<sup>10,29</sup>. Immune fluorescence and laser scanning confocal microscopy were performed as described<sup>10,29</sup> using a Leica SP2 microscope.

**Statistical analysis.** Statistical evaluation was performed as indicated within the legends using Sigma Stat software. *P* values are:  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ .

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 6 July; accepted 1 September 2009.

Published online 14 October 2009.

- Wekerle, H., Linington, C., Lassmann, H. & Meyermann, R. Cellular immune reactivity within the CNS. *Trends Neurosci.* **9**, 271–277 (1986).
- Ben-Nun, A., Wekerle, H. & Cohen, I. R. The rapid isolation of clonable antigen-specific T lymphocyte lines capable of mediating autoimmune encephalomyelitis. *Eur. J. Immunol.* **11**, 195–199 (1981).
- Flügel, A., Willem, M., Berkowicz, T. & Wekerle, H. Gene transfer into  $CD4^+$  T lymphocytes: Green fluorescent protein engineered, encephalitogenic T cells used to illuminate immune responses in the brain. *Nature Med.* **5**, 843–847 (1999).
- Odoardi, F., Kawakami, N., Klinkert, W. E. F., Wekerle, H. & Flügel, A. Blood-borne soluble protein antigen intensifies T cell activation in autoimmune CNS lesions and exacerbates clinical disease. *Proc. Natl Acad. Sci. USA* **104**, 18625–18630 (2007).
- Traugott, U. & Raine, C. S. Acute experimental allergic encephalomyelitis. Myelin basic protein-reactive T cells in the circulation and in meningeal infiltrates. *J. Neurol. Sci.* **42**, 331–336 (1979).
- Tsuchida, M. *et al.* Identification of  $CD4^+CD8^- \alpha\beta$  T cells in the subarachnoid space of rats with experimental autoimmune encephalomyelitis. A possible route by which effector cells invade the lesions. *Immunology* **81**, 420–427 (1994).
- Auffray, C. *et al.* Monitoring of blood vessels and tissues by a population of monocytes with patrolling behavior. *Science* **317**, 666–670 (2007).
- Phillipson, M. *et al.* Intraluminal crawling of neutrophils to emigration sites: a molecularly distinct process from adhesion in the recruitment cascade. *J. Exp. Med.* **203**, 2569–2575 (2006).
- Geissmann, F. *et al.* Intravascular immune surveillance by CXCR6<sup>+</sup> NKT cells patrolling liver sinusoids. *PLoS Biol.* **3**, e113 (2005).
- Flügel, A. *et al.* Migratory activity and functional changes of green fluorescent effector T cells before and during experimental autoimmune encephalomyelitis. *Immunity* **14**, 547–560 (2001).
- Vajkoczy, P., Laschinger, M. & Engelhardt, B.  $\alpha 4$ -integrin-VCAM binding mediates G protein independent capture of encephalitogenic T cell blasts to CNS white matter microvessels. *J. Clin. Invest.* **108**, 557–565 (2001).
- Kawakami, N. *et al.* Autoimmune  $CD4^+$  T cell memory: Lifelong persistence of encephalitogenic T cell clones in healthy immune repertoires. *J. Immunol.* **175**, 69–81 (2005).
- Wolf, K. *et al.* Compensation mechanism in tumor cell migration: mesenchymal–amoeboid transition after blocking of pericellular proteolysis. *J. Cell Biol.* **160**, 267–277 (2003).
- Yednock, T. A. *et al.* Prevention of experimental autoimmune encephalomyelitis by antibodies against  $\alpha 4\beta 1$  integrin. *Nature* **356**, 63–66 (1992).
- Schenkel, A. R., Mamdough, Z. & Muller, W. A. Locomotion of monocytes on endothelium is a critical step during extravasation. *Nature Immunol.* **5**, 393–400 (2004).
- Carman, C. V. *et al.* Transcellular diapedesis is initiated by invasive podosomes. *Immunity* **26**, 784–797 (2007).
- Walther, M. *et al.* Exogenous antigen containing perivascular phagocytes induce a non-encephalitogenic extravasation of primed lymphocytes. *J. Neuroimmunol.* **117**, 30–42 (2001).

18. Hickey, W. F., Vass, K. & Lassmann, H. Bone marrow derived elements in the central nervous system: an immunohistochemical and ultrastructural survey of rat chimeras. *J. Neuropathol. Exp. Neurol.* **51**, 246–256 (1992).
  19. Matyszak, M. K. & Perry, V. H. The potential role of dendritic cells in immune-mediated inflammatory diseases in the central nervous system. *Neuroscience* **74**, 599–608 (1996).
  20. McMenamin, P. G., Wealhall, R. J., Deverall, M., Cooper, S. J. & Griffin, B. Macrophages and dendritic cells in the rat meninges and choroid plexus: three-dimensional localisation by environmental scanning electron microscopy and confocal microscopy. *Cell Tissue Res.* **313**, 259–269 (2003).
  21. Kivisäkk, P. *et al.* Localizing central nervous system immune surveillance: Meningeal antigen-presenting cells activate T cells during experimental autoimmune encephalomyelitis. *Ann. Neurol.* **65**, 457–469 (2009).
  22. Reboldi, A. *et al.* C-C chemokine receptor 6-regulated entry of T<sub>H</sub>-17 cells into the CNS through the choroid plexus is required for the initiation of EAE. *Nature Immunol.* **10**, 514–523 (2009).
  23. Becher, B., Bechmann, I. & Greter, M. Antigen presentation in autoimmunity and CNS inflammation: how T lymphocytes recognize the brain. *J. Mol. Med.* **84**, 532–543 (2006).
  24. Bechmann, I., Galea, I. & Perry, V. H. What is the blood-brain barrier (not)? *Trends Immunol.* **28**, 5–11 (2007).
  25. Hinrichs, D. J., Wegmann, K. W. & Dietsch, G. N. Transfer of experimental allergic encephalomyelitis to bone marrow chimeras: endothelial cells are not the restricting element. *J. Exp. Med.* **166**, 1906–1911 (1987).
  26. Hickey, W. F. & Kimura, H. Perivascular microglial cells of the CNS are bone-marrow derived and present antigen *in vivo*. *Science* **239**, 290–292 (1988).
  27. Lois, C., Hong, I. J., Pease, S., Brown, E. J. & Baltimore, D. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* **295**, 868–872 (2002).
  28. Kawakami, N. *et al.* Live imaging of effector cell trafficking and autoantigen recognition within the unfolding autoimmune encephalomyelitis lesion. *J. Exp. Med.* **201**, 1805–1814 (2005).
  29. Kawakami, N. *et al.* The activation status of neuroantigen-specific T cells in the target organ determines the clinical outcome of autoimmune encephalomyelitis. *J. Exp. Med.* **199**, 185–197 (2004).
- Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).
- Acknowledgements** We thank S. Kosin, I. Haarmann, C. Federle, T. H. Ngugen and K. Nispel for technical assistance. We thank M. Krumbholz for help in statistical analysis, M. Hübener for support in 3D analysis software, and K. Dornmair, E. Meinel and E. Lütjen-Drecoll for critical comments on the paper. We are grateful to V. Staiger for providing software and R. Schorner for help with the artwork. This work was supported by the Deutsche Forschungsgemeinschaft (SFB455 project A8 and SFB571 project C6) and the Hertie foundation (Hertie no.1.01.1/04/010).
- Author Contributions** I.B., N.K. and C.S. performed most of the imaging experiments. N.K., F.O. and D.M. performed the experiments testing the antigen presentation capacity of meningeal/perivascular macrophages. F.O. performed the transcriptome analyses, EAE studies and cytofluorometric characterizations. W.E.F.K. and J.W.E. performed cell sorting. C.F.-K. contributed to the immunohistochemical studies. T.B.I. provided a monoclonal antibody. A.F. coordinated the experimental work. A.F. and H.W. designed the study and wrote the manuscript with inputs from co-authors.
- Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to A.F. ([Fluegel@med.uni-goettingen.de](mailto:Fluegel@med.uni-goettingen.de)) or H.W. ([hwekerle@neuro.mpg.de](mailto:hwekerle@neuro.mpg.de)).



## METHODS

**Animals and antigens.** Lewis rats and SJL/J mice were obtained from the animal facility of the Max-Planck Institute, and were kept under standardized conditions. GFP<sup>+</sup> rats were provided by D. Baltimore<sup>27</sup>. GFP<sup>+</sup> mice were from M. Okabe<sup>30</sup> and were back-crossed to SJL/J background for more than 10 generations. MBP was prepared from guinea-pig brain<sup>31</sup>. Ovalbumin was from Sigma, PLP<sub>139–151</sub> (proteolipid protein peptide) from Biotrend, FITC-, Texas-red- and tetramethylrhodamine-conjugated dextran (molecular mass 70 kDa) came from Invitrogen.

**T cells and EAE induction.** CD4<sup>+</sup> T<sub>MBP-GFP/RFP</sub> and T<sub>OVA-GFP/RFP</sub> cell lines were generated and tested for phenotype, cytokine profile, and antigen specificity as described<sup>33,32</sup>. Adoptive transfer EAE was induced by intravenous injection of  $5 \times 10^6$  T<sub>MBP-GFP/RFP</sub> cells. Weight and clinical scores were scored daily (score 0 = no disease; 1 = flaccid tail; 2 = gait disturbance; 3 = complete hind limb paralysis; 4 = tetraparesis; and 5 = death).

For generation of murine T cell lines, SJL GFP<sup>+</sup> mice were immunized subcutaneously with 100 µg of PLP<sub>139–151</sub> in complete Freund's adjuvant (containing 4 mg ml<sup>-1</sup> of heat-killed *Mycobacterium tuberculosis* H37Ra, Difco). Ten days later draining lymph nodes were recovered and cultured in presence of the antigen as described<sup>33</sup>. EAE was induced by transfer of  $5 \times 10^6$  T<sub>PLP-GFP</sub> cells per animal.

**Active EAE induction in memory animals.** Memory animals were established by intraperitoneal transfer of T<sub>MBP-GFP</sub> cells ( $2 \times 10^6$  cells per animal, 5 days after stimulation with MBP-pulsed antigen-presenting cells, in 0.5 ml of Eagle's HEPES medium) into newborn Lewis rats as described<sup>12</sup>. EAE was induced in 8–10-week-old T<sub>MBP-GFP</sub> memory animals by active immunization with MBP (200 µg per animal) in complete Freund's adjuvant.

**Intravital imaging: surgical procedures.** Animals were anaesthetized by intramuscular injection of fentanyl/midazolam/medetomidin (5 µg kg<sup>-1</sup>, 2 mg kg<sup>-1</sup> and 150 µg kg<sup>-1</sup>, respectively), orotracheally intubated and ventilated with 1.5–2% isoflurane. A spinal cord window for imaging was prepared at level Th12/L1. After midline skin incision of 2–3 cm the paravertebral musculature was detached from the spine and a laminectomy on one spine disc was performed. The dura was left intact whenever possible. To exclude artefacts by the preparative disclosure of the spinal cord tissue, we also performed imaging of the tissue between adjacent lumbar vertebral arches, leaving bone structures completely intact. Both preparation techniques gave comparable results (not shown).

For peripheral nerve imaging a skin flap (diameter ~1 cm) was removed from dorsal hind limbs. The veins of the tibial and saphenous nerves were exposed by careful dislocation of the superposed muscle layer. The ear was carefully separated in a dorsal and a ventral half. Blood vessels of the ventral half were exposed.

Animals were stabilized in a custom-made microscope stage and body temperature was regulated by a heated pad (37.5 °C). Electrocardiograms were recorded during the imaging. In some experiments, antibodies were infused by a tail vein catheter (see later).

**Intravital imaging: technical equipment.** Time-lapse two-photon laser-scanning microscopy was performed using a Leica SP2 confocal microscope (Leica Microsystems) combined with a 10W-Millenia/Tsunami laser (Spectra Physics)<sup>4,32</sup>. The excitation wavelength was tuned to 880 or 935 nm and routed through an Olympus  $\times 20$  water immersion objective (NA 0.95). Typically areas of 600  $\times$  600 µm width were scanned and 30–50 µm z-stacks were acquired (3–5 µm z-step). The acquisition rate was approximately 1.3 s per plane, that is, 32 s interval time. Maximum penetration depths reached up to 200 µm below the meningeal blood vessels. The images were averaged two to four times. Fluorescent signals were detected using non-descanned detectors (Hamamatsu) equipped with 525/50 nm and 630/69 nm band-pass filters (Semrock).

**Intravital imaging: image analysis.** Time-lapse images were processed and analysed by Leica LCS software (Leica), Imaris (Bitplane) and ImageJ (NIH). To obtain two-dimensional movies, a Gaussian blur filter was applied and maximum intensity z-projections were made. Three-dimensional reconstructions were generated using Imaris software. To determine T-cell velocity, four-dimensional data were generated by the Imaris Track module and revised manually. Motility parameters and trajectories were calculated from the obtained position coordinates using Excel (Microsoft).

**Intravital imaging: analysis of T cell motility.** The average crawling time has been determined analysing all T cells in continuous recordings of up to 1 h. This analysis included cells present at the beginning and end of imaging as well as tracks exceeding the observation time. For the correlation of crawling versus rolling T cells, the average number of crawling cells per time point was considered. Rolling T cells were defined as cells appearing as single or several round-shaped dots or cells moving solely in direction of the blood flow with  $>50 \mu\text{m min}^{-1}$  instantaneous velocity. Analysis was performed by quantifying the absolute number of rolling events within 30 min recording time. Owing to limitations in

temporal resolution of the two-photon scanning equipment used, fast rolling events might have escaped our analysis. The meandering index is defined as the sum of total displacement divided by path length of a T cell.

**Intravital imaging: cell and vessel labelling.** Texas-red- or FITC-conjugated dextran (40 µg per rat; 70 kDa) was stereotactically injected into the cisterna magna of anaesthetized animals (10 mg kg<sup>-1</sup> xylazine combined with 50 mg kg<sup>-1</sup> ketamine) 1 day before transfer of T cells or *ex vivo*-isolation of the fluorescently labelled antigen-presenting cells. Blood vessels were filled by intravenous infusion of fluorescent dextran or peridinin chlorophyll protein complex-conjugated anti-human CD4 antibody.

**Immunostaining and confocal microscopy.** For analyses of T<sub>MBP-GFP</sub> cell infiltration kinetics, animals were perfused with 4% paraformaldehyde at graded time points. For immunofluorescence spinal cord cryosections were incubated with 5% BSA in PBS for 2 h at room temperature before addition of primary antibodies: anti-smooth muscle actin-Cy3-conjugate (Sigma, 1:100) or anti-fibronectin (Dako Cytomation, 1:100). Secondary antibody was Alexa 555-conjugated anti-rabbit antisera (Jackson ImmunoResearch, 1:500). Stained samples were embedded in mounting medium (MP Biomedicals). Images were taken with a Leica SP2 confocal laser scanning microscope equipped with a  $\times 40$  oil immersion objective (NA 1.25) or a  $\times 20$  air immersion objective (NA 0.70).

**Monoclonal antibody treatments.** To block integrin-mediated binding, antibodies neutralizing CD11a (integrin  $\alpha\text{L}$ , anti-LFA-1, clone WT.1, Serotec) and anti-CD49d (integrin  $\alpha\text{4}$ , anti-VLA-4, clone TA-2)<sup>34,35</sup> were infused intravenously at 4 mg kg<sup>-1</sup> before or during intravital imaging. Flow cytometric analysis after intravital recording was performed to control for saturated binding of the respective antibodies (not shown).

**Cell isolation, flow cytometry and FACS.** T<sub>MBP-GFP</sub> cells were isolated from organs and blood as follows<sup>10,32</sup>. After lysis of the splenic erythrocytes by osmotic shock, the remaining white cells were cultured at 37 °C for 1 h in DMEM plus 10% FCS to remove the adherent cells. Lymphocytes were purified from blood (obtained by heart puncture) by density gradient centrifugation (Nycomed). Meninges were carefully dissected from parenchymal spinal cord and mechanically dissociated by trituration and passage through a 70-µm cell strainer. T cells from CNS parenchyma were purified by Percoll gradient centrifugation (GE Healthcare). Immediately after cell preparation, T<sub>GFP</sub> cells were sorted in PBS containing 2% glucose at 4 °C using FACS Vantage (BD) or MoFlo (Dako). For cell-surface marker staining<sup>29</sup> we used the following antibodies: CD4 (W3/25), CD8 (OX-8),  $\alpha\text{B2CR}$  (R73), CD45RA/B (OX-33, Cedarlane Laboratories), MHC I (OX-18), MHC II (OX-6), CD106 (VCAM1), RECA1 (HIS52), CD11a (WT.1), CD11b (OX-42), CD11c (8A2, integrin  $\alpha\text{X}$  chain), CD25 (OX-39), CD134 (OX-40). Mouse IgG1 $\kappa$  (MOPC 31 C, Sigma) served as an isotype control. Allophycocyanin-labelled anti-mouse antibody (Invitrogen) was used as secondary antibody. All antibodies were purchased from Serotec, unless stated otherwise. For intracellular staining, the isolated T<sub>MBP-GFP</sub> cells were incubated with brefeldin (5 µg ml<sup>-1</sup>, Sigma) for 4 h, fixed and permeabilized using Cytofix/Cytoperm buffer (BD) and then incubated with anti-cytokines antibodies (mouse anti-rat IFN- $\gamma$  (DB-1, Invitrogen) and rat phycoerythrin (PE)-conjugated anti-mouse IL17-PE, BD). Mouse IgG1 $\kappa$  (MOPC 31 C) and PE rat IgG1 were the isotype control. Cytofluorometric analysis was performed with FACS-Calibur operated by Cell Quest software (BD). All procedures were performed at 4 °C.

**In vivo proliferation and cell cycle analysis.** BrdU (120 mg kg<sup>-1</sup> body weight; BD) was injected intraperitoneally 2 h before organ recovery. This interval between BrdU injection and collection was chosen to minimize cell trafficking and to allow BrdU to enter the CNS parenchyma<sup>36</sup>. BrdU-containing T<sub>MBP-GFP</sub> cells were isolated from spleen, blood, meninges and CNS parenchyma sequentially after transfer and stained using BrdU Flow kit (BD) according to the manufacturer's instructions. In brief, the cells were fixed and permeabilized with BD Cytofix/Cytoperm buffer, treated with DNase to expose BrdU epitopes, stained with allophycocyanin-labelled anti-BrdU antibodies and resuspended in 7-amino-actinomycin D (7-AAD) solution for cell cycle analysis. The cells were then analysed by flow cytometry.

**Generation of GFP<sup>+</sup> bone marrow chimaeras.** Bone marrow cells were prepared from femurs of 3–4-months-old GFP<sup>+</sup> Lewis rats<sup>27</sup>. Bone marrow cells suspended in DMEM were transferred intraperitoneally into irradiated (9 Gy) wild-type Lewis rats ( $6 \times 10^8$  cells per recipient). Reconstitution of the recipient bone marrow was checked by cytofluorometry ( $>90\%$  replacement, not shown). Animals were used at least 10 weeks after transplantation to guarantee maximal replacement of perivascular and meningeal phagocytes<sup>18</sup>.

**Intrathecal transfer of meningeal phagocytes.** Fluorochrome-labelled antigen-presenting cells were isolated by FACS and pulsed with OVA or MBP (10 µg ml<sup>-1</sup>) in DMEM plus 1% rat serum at 37 °C for 1 h. The loaded cells were washed three times, resuspended in endotoxin-free buffer and re-transferred intrathecally ( $1 \times 10^7$  total meningeal cells or  $1 \times 10^4$  sorted fluorescent phagocytes) into hosts

that had received  $1 \times 10^7$   $T_{OVA-GFP}$  cells 3 days earlier. Animals were analysed 24 h after antigen-presenting cell transfer.

**In vitro T cell stimulation by ex vivo-isolated meningeal antigen-presenting cells or meningeal explants.** Meningeal antigen-presenting cells, labelled by intrathecal injection of FITC-conjugated dextran for 24 h, were isolated from spinal cord homogenates by FACS. The antigen-presenting cells were irradiated and  $1 \times 10^4$  cells were co-cultured with equal numbers of resting  $T_{MBP-GFP}$  or  $T_{OVA-GFP}$  cells in 96-well plates (in DMEM plus 1% rat serum). As a positive control, irradiated thymocytes containing 'professional' antigen-presenting cells (thymic dendritic cells) with or without specific antigens were used. For transcriptome analysis, cells were collected after 24 h in TRIzol (Invitrogen) and stored at  $-80^\circ\text{C}$  until mRNA preparation. To determine proliferation, co-cultures were measured after 48 h with  $[^3\text{H}]-2'$ -deoxy-thymidine ( $[^3\text{H}]\text{dT}$ ; 2 Ci mmol $^{-1}$ ; Amersham), and isotope incorporation was determined 16 h later in a  $\beta$ -counter (Coulter). For intracellular IFN- $\gamma$  staining, meningeal explants or single cell homogenates were used. Meningeal explants were prepared from naive Lewis rat and co-cultured with  $1 \times 10^4$   $T_{MBP-GFP}$  and  $1 \times 10^4$   $T_{OVA-RFP}$  cells in 96-well plates for 36 h. Single cell homogenates were prepared from naive or OVA intrathecally treated animals. After irradiation (50 Gy),  $1 \times 10^6$  cells were co-cultured with  $2 \times 10^4$   $T_{MBP-GFP}$  or  $T_{OVA-GFP}$  cells in a 96-well plate (in DMEM plus 1% rat serum) for 16 h. IFN- $\gamma$  staining was performed as described earlier.

**Quantitative PCR.** For transcriptome profiling, the sorted  $T_{MBP-GFP}$  cells were sorted in TRIzol (Invitrogen). Total mRNA preparation and cDNA reverse transcription were performed using standard protocols (Sigma-Aldrich and Invitrogen). TaqMan analysis<sup>10</sup> used ABI Prism 5700 Sequence Detector 'TaqMan' (PE Applied Biosystems).  $\beta$ -actin served as a housekeeping gene standard. Data were obtained by at least duplicate measurements. The  $C_T$  value of the individual measurements did not exceed 0.5 amplification cycles. The rat primers and probes used for detection of activation markers and chemokines receptors/ligands were recently described<sup>32</sup>. The other primers/probes used are listed here: *CD3*: forward: 5'-CAAAGAACTAACATGGAGCAGGG-3', backward: 5'-CTTTTGGCTGGCCATGGT-3', probe: 5'-AGGTTTGGCTGGCCTCTTCCTGGTG-3'; *iNOS* (also known as *Nos2*): forward: 5'-TTCCCATCGCTCCGCTG-3', backward: 5'-CCGAGCTGTAGCACTGCA-3', probe: 5'-AACACAGTAATGGCCGACCTGATGTGC-3'; *Tnf*: forward: 5'-TCGAGTGACAAGCCGTAGC-3', backward: 5'-CTCAGCCACTCCAGCTGCTC-3', probe: 5'-CGTCGTAGCAAAACCAAGCAGA-3'; *Il10*: forward: 5'-GAAGACCCTCTGATACAGCTGC-3', backward: 5'-TGCTCCACTGCCTTGTCTTTT-3', probe: 5'-CGCTGTCATCGATTTCTCCCTGTGA-3', probe: 5'-CGCTGTCATCGATTTCTCCCTGTGA-3'; *Tgfb*: forward: 5'-CCCTGCCCCATATTTTGA-3', backward: 5'-ACGGTGATGCGGAAGCAC-3', probe: 5'-CACACAGTACAGCAAGGTCCTTGCCCT-3'; *Il4*: forward: 5'-CGGTGAAGTGGAGAACTCTGTAGA-3', backward: 5'-TCAGTGTGTGAGCGTGGACTC-3', probe: 5'-CGGTCTGAAGTCACTGAGAAGCTGCACC-3'. *Itgb2* (*CD18*): forward: 5'-CTGCTACTCATCTAGCTGGACTACTC-3', backward: 5'-GCAGTTGCTGACTTGTACTTGGT-3', probe: 5'-TCCTGGGATCTGCCCTGTCC-3'; *Vla-4* (also known as *Itga4*): forward: 5'-GAAGGAAGAGTGTTCGTGTACATCA-3', backward: 5'-CGACGAGCACTTTTCCATT-3', probe: 5'-CTCTGGCATGGGAGCTGTGATGGTT-3'. *SIP1* (also known as *S1pr1*): forward: 5'-GATCGCGCGCGGTGTAGAC-3', backward: 5'-TTTCCTTGGCTGGAGAGG-3', probe: 5'-TTGAGCGAGGCTGCTGTTTCTC-3'; *Mmp9*: forward: 5'-CACGACAGCTGACTACGACACA-3', backward: 5'-ATGGAAATACGAGGGTTTGC-3', probe: 5'-TATGGTTTCTGCCCCAGTGAGAATC-3'; *Mmp14*: forward: 5'-CTGAACCTCTCATTGAAGCG-3', backward: 5'-CTGCATCAATCTTGTG

CGTGG-3', probe: 5'-ATTTCCCCGGAAGAAGTAGGTCTTC-3'; perforin: forward: 5'-GGTGGAGTGGAGGCTTTTGTG-3', backward: 5'-CCGAGAAGGCCATCAGG-3', probe: 5'-CCAGGCGAAAAGTGTACATGCGACACT-3'; FasL: forward: 5'-ACCAACCACAGCCTTAGAGTATCATC-3'; backward: 5'-CTGTAAAGTGGGCCACACTCCTT-3', probe: 5'-AAAGCAAATAGCCAACCCAGCACACC-3'; *Bcl2*: forward: 5'-TGAACCGGCATCTGCACA-3', backward: 5'-CAGAGGTGCGATGCTGGG-3', probe: 5'-AACGGAGGCTGGGATGCTTTGTG-3'; *Bcl-xl* (also known as *Bcl2l1*): forward: 5'-GGTGAGTCGGATTGCAAGTTG-3', backward: 5'-GTAGAGATCCACAAAAGTGTCCCAG-3', probe: 5'-CCTGAATGACCACCTAGAGCCTTGGATCC-3'; *Klf2*: forward: 5'-GCACCTAAAGGCGCATCTG-3', backward: 5'-AGCGCGCAACTTCCA-3', probe: 5'-AGGTGAGAAGCCTTATC-3'; *Egr2*: forward: 5'-GTGCCAGCTGCTACCCAGAAG-3', backward: 5'-GGCTGTGGTTGAAGCTGGAG-3', probe: 5'-TTGTGAGTGGCGCATCTTGC-3'; *Egr3*: forward: 5'-CAACGACATGGGCTCCATTC-3', backward: 5'-GGCCTTGATGGTCTCCAGAG-3', probe: 5'-CCTTCCAGGGCATGGACC-3'; *Tob-1*: forward: 5'-TTGACGATGTTCGTGGCAAT-3', backward: 5'-TTCTCACCGATTTGGTAGGAA-3', probe: 5'-TGCCCGCAGGATCTGAGCGTG-3'; cyclin B1: forward: 5'-TTGCTTAGCCCTGAAAATCCTT-3', backward: 5'-TCCTCCGTGTGGGACAGGTA-3', probe: 5'-AACGGTGAATGGACAGGA-3'; cyclin B2: forward: 5'-GACTCTGTACATGTGCTATGCTATCA-3', backward: 5'-CCCAACCAAGTTCAGCTTCT-3', probe: 5'-TTCCTTCAGGCTCAGCCA-3'; cyclin D3: forward: 5'-GGAATCAAGCCGCACATG-3', backward: 5'-TTCATAGCCAGAGGGAAGACATC-3', probe: 5'-ATGCTGGAGGTGTGTGAG-3'; cyclin E: forward: 5'-CATTTTCTCGTTGGAGTTGAT-3', backward: 5'-GGAACCATCCACTTGACACATT-3', probe: 5'-AAGGTCTCAGGTTATCAGTG-3'; *p21* (also known as *Cdkn1a*): forward: 5'-CACTCCAAGCGCCGATTG-3', backward: 5'-CCTGACCCACAGCAGAAGAA-3', probe: 5'-AAAGCCCTGAAGTGCCAC-3'; *Rora*: forward: 5'-GAAGTATGCTAGCCCCGATGTC-3', backward: 5'-CTTCCAAATTCAAACACAAAGCT-3', probe: 5'-TCAAATCCTTAGGTTGTGAAGACTC-3'; *Rorg*: forward: 5'-GACAGGGCCCCACAGAGA-3', backward: 5'-TTTGTGAGGTGTGGGTCTTCTTT-3', probe: 5'-CGAACATCTCGGGAGTTGCTGGCT-3'; *T-bet* (also known as *Tbx21*): forward: 5'-CCAACAATGTGACCCAGATGAT-3', backward: 5'-CTGGCTCACCGTCATTCA-3', probe: 5'-TCCTGCAGTCCCTCCATAAGTACCAGCC-3'.

30. Hadjantonakis, A.-K., Gertsenstein, M., Ikawa, M., Okabe, M. & Nagy, A. Generating green fluorescent mice by germline transmission of green fluorescent ES cells. *Mech. Dev.* **76**, 79–90 (1998).
31. Eylar, E. H., Kniskern, P. J. & Jackson, J. J. Myelin basic proteins. *Methods Enzymol.* **32**, 323–341 (1974).
32. Odoardi, F. *et al.* Instant effect of soluble antigen on effector T cells in peripheral immune organs during immunotherapy of autoimmune encephalomyelitis. *Proc. Natl Acad. Sci. USA* **104**, 920–925 (2007).
33. Kuchroo, V. K. *et al.* Experimental allergic encephalomyelitis mediated by cloned T cells specific for a synthetic peptide of myelin proteolipid protein. Fine specificity and T cell receptor V $\beta$  usage. *J. Immunol.* **148**, 3776–3782 (1992).
34. Tamatani, T., Kotani, M. & Miyasaka, M. Characterization of the rat leukocyte integrin, CD11/CD18, by the use of LFA-1 subunit specific monoclonal antibodies. *Eur. J. Immunol.* **21**, 627–633 (1991).
35. Issekutz, T. B. Dual inhibition of VLA-4 and LFA-1 maximally inhibits cutaneous delayed-type hypersensitivity-induced inflammation. *Am. J. Pathol.* **143**, 1286–1293 (1993).
36. Kornack, D. R. & Rakic, P. Changes in cell-cycle kinetics during the development and evolution of primate neocortex. *Proc. Natl Acad. Sci. USA* **95**, 1242–1246 (1998).

# HMGB proteins function as universal sentinels for nucleic-acid-mediated innate immune responses

Hideyuki Yanai<sup>1\*</sup>, Tatsuma Ban<sup>1\*</sup>, ZhiChao Wang<sup>1\*</sup>, Myoung Kwon Choi<sup>1</sup>, Takeshi Kawamura<sup>2</sup>, Hideo Negishi<sup>1</sup>, Makoto Nakasato<sup>1</sup>, Yan Lu<sup>1</sup>, Sho Hangai<sup>1</sup>, Ryuji Koshiba<sup>1</sup>, David Savitsky<sup>1</sup>, Lorenza Ronfani<sup>3</sup>, Shizuo Akira<sup>4</sup>, Marco E. Bianchi<sup>3</sup>, Kenya Honda<sup>1†</sup>, Tomohiko Tamura<sup>1</sup>, Tatsuhiko Kodama<sup>2</sup> & Tadatsugu Taniguchi<sup>1</sup>

The activation of innate immune responses by nucleic acids is crucial to protective and pathological immunities and is mediated by the transmembrane Toll-like receptors (TLRs) and cytosolic receptors<sup>1,2</sup>. However, it remains unknown whether a mechanism exists that integrates these nucleic-acid-sensing systems. Here we show that high-mobility group box (HMGB) proteins 1, 2 and 3 function as universal sentinels for nucleic acids. HMGBs bind to all immunogenic nucleic acids examined with a correlation between affinity and immunogenic potential. *Hmgb1*<sup>-/-</sup> and *Hmgb2*<sup>-/-</sup> mouse cells are defective in type-I interferon and inflammatory cytokine induction by DNA or RNA targeted to activate the cytosolic nucleic-acid-sensing receptors; cells in which the expression of all three HMGBs is suppressed show a more profound defect, accompanied by impaired activation of the transcription factors interferon regulatory factor 3 (IRF3) and nuclear factor (NF)- $\kappa$ B. The absence of HMGBs also severely impairs the activation of TLR3, TLR7 and TLR9 by their cognate nucleic acids. Our results therefore indicate a hierarchy in the nucleic-acid-mediated activation of immune responses, wherein the selective activation of nucleic-acid-sensing receptors is contingent on the more promiscuous sensing of nucleic acids by HMGBs. These findings may have implications for understanding the evolution of the innate immune system and for the treatment of immunological disorders.

During microbial infection or tissue damage, DNA and RNA potentially activate the innate and subsequent adaptive immune responses<sup>1,2</sup>. In mammals, TLR3, TLR7 and TLR9 recognize, respectively, double-stranded RNA, single-stranded and short double-stranded RNAs, and hypomethylated DNA<sup>1-3</sup>, whereas the RIG-I-like receptors (RLRs), namely retinoic acid-inducible gene I (RIG-I) and melanoma differentiation-associated gene-5 (MDA5) are best known as RNA-sensing receptors in the cytosol<sup>4,5</sup>. In addition, cytosolic DNA-sensing receptors, which include DNA-dependent activator of IRFs (DAI) and absent in melanoma 2 (AIM2), also trigger the innate and adaptive immune systems<sup>6-8</sup>. It has recently been shown that RLRs also participate in the cytosolic DNA-sensing system<sup>3,9-11</sup>. The hallmark of innate immune responses activated by these receptors is the induction of type-I interferons (IFNs), proinflammatory cytokines and chemokines<sup>1</sup>, except that by AIM2, which is a critical component of the inflammasome that typically promotes the secretion of interleukin (IL)-1 $\beta$  (ref. 7). So far, no universal or shared mechanism of action for these nucleic-acid-receptor classes is presumed to operate in their activation.

To gain new insights into the nature of the cytosolic DNA-sensing systems, we performed an unbiased biochemical screen to identify

proteins involved in DNA recognition on the basis of their direct binding to a B-form DNA, poly(dA-dT)•poly(dT-dA) (B-DNA; ref. 12); the most prominently recovered proteins were HMGB1, HMGB2 and HMGB3 (Supplementary Fig. 1a). HMGB proteins are highly expressed in the nucleus, where they regulate chromatin structure and transcription, but they are also present in the cytosol and in extracellular fluids<sup>13,14</sup>. HMGB1 was recently shown to participate in the activation of several immune receptors, including TLRs<sup>13,15-18</sup>.

To examine the direct interaction of HMGBs with nucleic acids, recombinant HMGBs were purified and assayed *in vitro* for their binding to biotin-conjugated B-DNA. Both HMGB1 and HMGB2 were precipitated by immobilized B-DNA, which was inhibited in a dose-dependent manner by non-conjugated B-DNA (Supplementary Fig. 1b). HMGB1-B-DNA binding was, however, quite inefficiently inhibited by the addition of calf thymus-derived or bacteria-derived DNA, each of which only weakly activates the cytosolic-sensing pathways<sup>8,12</sup> (Supplementary Fig. 1b). The addition of double-stranded RNA (polyinosinic-polycytidylic acid; poly(I:C)) or single-stranded RNA (poly(U)), but not imiquimod (R837), a non-nucleic-acid agonist for TLR7, efficiently inhibited HMGB1-B-DNA binding (Supplementary Fig. 1b). In contrast, HMGB2-B-DNA binding was not affected by either poly(I:C) or poly(U) (Supplementary Fig. 1b). Thus, HMGB1 binds to these immunogenic RNAs but HMGB2 does not. Furthermore, a TLR9 agonist CpG-B oligodeoxynucleotide (ODN; refs 19, 20) and antagonist base-free phosphorothioate deoxyribose homopolymer (PS; ref. 21) were the most potent in inhibiting HMGB1-B-DNA binding, which is consistent with a previous report<sup>16</sup> (Supplementary Fig. 1b). In contrast, a weak TLR9 agonist base-free natural phosphodiester deoxyribose homopolymer (PD; ref. 21) showed little, if any, inhibition (Supplementary Fig. 1b). Binding of HMGB1 to biotin-conjugated poly(U), a TLR7 agonist, was also strongly inhibited by the addition of free CpG-B ODN and PS, but not by R837 (Supplementary Fig. 1c; see below). Finally, we also found that HMGB3, expressed in certain cell types<sup>22,23</sup>, binds both DNA and RNA (Supplementary Fig. 1d). These results indicate a correlation between the affinity of a type of nucleic acid to HMGB and its immunogenicity.

To study the contribution of HMGBs to the nucleic-acid-mediated activation of innate immune responses, we first examined cells derived from gene-targeted mice for HMGB1 or HMGB2. Mouse embryonic fibroblasts (MEFs) from HMGB1-deficient mice (*Hmgb1*<sup>-/-</sup>) showed a significant defect in messenger RNA induction for type-I IFNs, IL-6 and RANTES in response to cytosolically delivered B-DNA or poly(I:C) at all doses examined, whereas the response to lipopolysaccharide (LPS)

<sup>1</sup>Department of Immunology, Graduate School of Medicine and Faculty of Medicine, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan. <sup>2</sup>Laboratory for System Biology and Medicine, RCAST, University of Tokyo, Komaba 4-6-1, Meguro-ku, Tokyo 153-8904, Japan. <sup>3</sup>Faculty of Medicine, San Raffaele University, via Olgettina 58, 20132 Milan, Italy. <sup>4</sup>Laboratory of Host Defense, WPI Immunology Frontier Research Center, Osaka University, Yamada-oka 3-1, Suita, Osaka 565-0871, Japan. <sup>†</sup>Present address: Department of Microbiology and Immunology, Graduate School of Medicine, Osaka University, Yamada-oka 2-2, Suita, Osaka 565-0871, Japan.

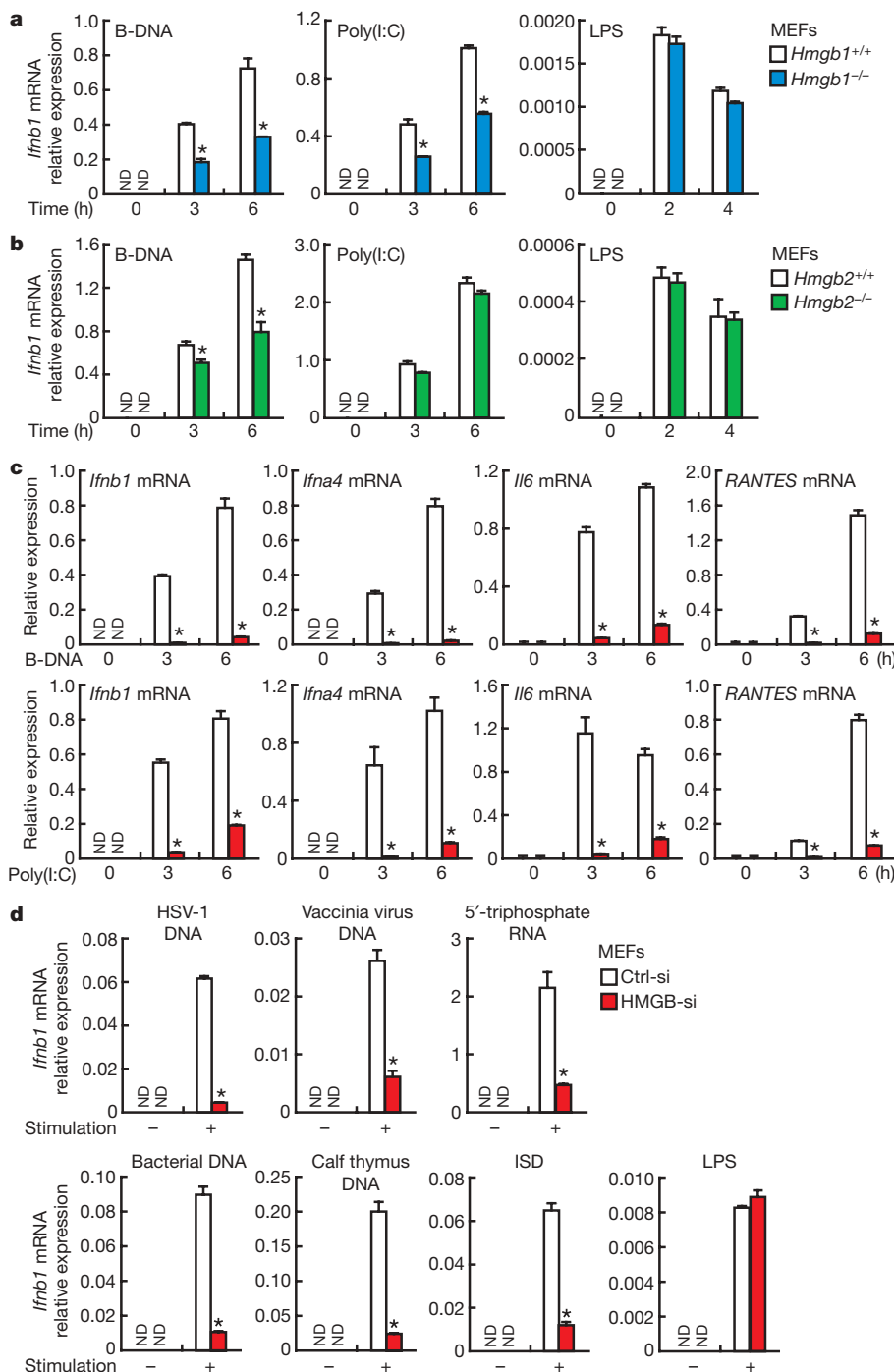
\*These authors contributed equally to this work.



remained unaffected (Fig. 1a and Supplementary Fig. 2a, b). Similar results were obtained with conventional dendritic cells (cDCs) differentiated by culturing *Hmgb1*<sup>-/-</sup> fetal liver with granulocyte-macrophage colony-stimulating factor (GM-CSF) (Supplementary Fig. 2c). Cytokine gene induction by *Hmgb2*<sup>-/-</sup> MEFs was defective when stimulated by B-DNA and not poly(I:C), which is consistent with the interaction of HMGB2 with DNA only (Fig. 1b and Supplementary Figs 1b and 2d). Accordingly, in *Hmgb1*<sup>-/-</sup> MEFs expressing a small interfering RNA (siRNA) that specifically targets HMGB2, type-I IFN gene induction on stimulation with B-DNA, but not with poly(I:C), was decreased further; the residual response is presumed to have been due to HMGB3 (Supplementary Fig. 2e, f).

We then examined MEFs in which the expression of all three HMGB proteins was suppressed by use of pan-HMGB siRNA vector (Supplementary Fig. 2g, h). As shown in Fig. 1c, the mRNA induction for type-I IFNs and other cytokines on stimulation with B-DNA

or poly(I:C) was inhibited more strongly in MEFs expressing the pan-HMGB siRNA (HMGB-si-MEFs) than in the *Hmgb1*<sup>-/-</sup> and *Hmgb2*<sup>-/-</sup> cells tested above. Similar results were obtained on stimulation of HMGB-si-MEFs with IFN-stimulatory DNA (ISD; refs 2, 24), viral or bacterial DNA, and single-stranded RNA bearing 5'-triphosphates<sup>25,26</sup> (Fig. 1d and Supplementary Fig. 2i). In contrast, the induction of cytokine genes remained unaffected in HMGB-si-MEFs stimulated by LPS, indicating that the inhibition is selective to stimulations with nucleic acids (Fig. 1d and Supplementary Fig. 2i-k). Moreover, mRNA induction of the genes activated by a variety of cytokines occurred normally in these cells, indicating that gene transcription is not generally affected by the absence of HMGBs (Supplementary Fig. 2l, m). Essentially the same observations were also made in RAW264.7 macrophage and NIH/3T3 fibroblast cell lines (Supplementary Fig. 2n-q). In macrophages, cytosolic DNA induces the formation of the inflammasome by the activation of AIM2, which

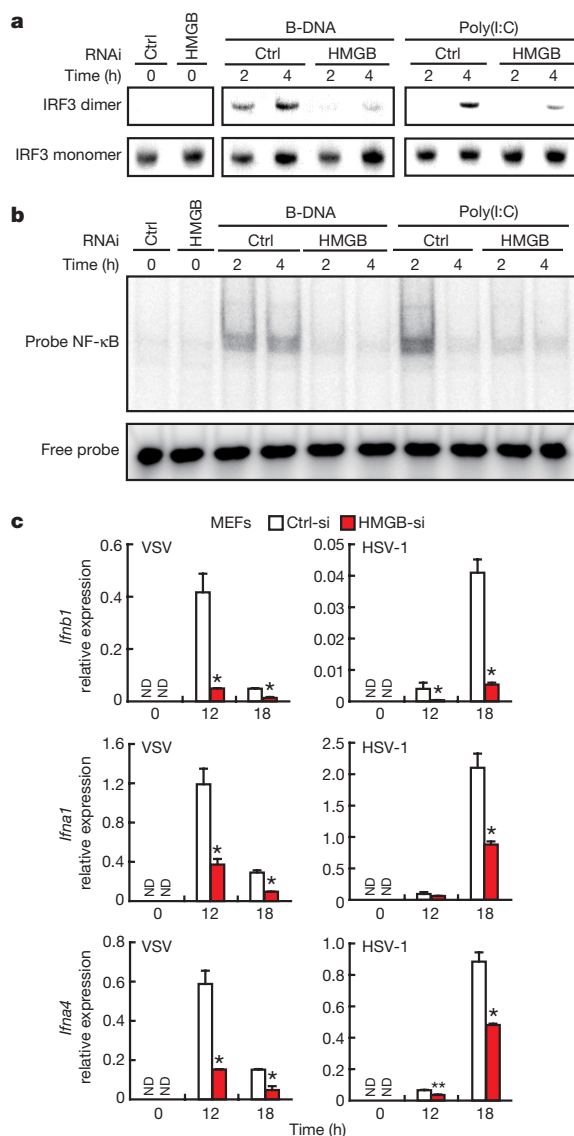


**Figure 1 | The contribution of HMGBs to cytosolic DNA-mediated or RNA-mediated activation of innate immune responses.**

**a**, Hypoinduction of IFN- $\beta$  in *Hmgb1*<sup>-/-</sup> cells on cytosolic delivery of DNA or RNA. *Hmgb1*<sup>+/+</sup> or *Hmgb1*<sup>-/-</sup> MEFs were stimulated with B-DNA or poly(I:C) for 6 h, or stimulated with LPS (200 ng ml<sup>-1</sup>) for 2 h. The induction levels of *Ifnb1* mRNA were determined by quantitative real-time RT-PCR (qRT-PCR). Asterisk, *P* < 0.01 compared with *Hmgb1*<sup>+/+</sup> MEFs. Data in all panels are presented as means and s.d. (*n* = 3). ND, not detected. **b**, Hypoinduction of IFN- $\beta$  in *Hmgb2*<sup>-/-</sup> MEFs on cytosolic delivery of DNA but not RNA. Asterisk, *P* < 0.001 compared with *Hmgb2*<sup>+/+</sup> MEFs. **c**, Impaired innate immune responses to cytosolic nucleic acids in MEFs deficient in all HMGBs. MEFs transduced with retrovirus expressing siRNA targeting all HMGBs (HMGB-si) or control siRNA (Ctrl-si) were lipofected with B-DNA or poly(I:C), and then mRNA expression levels of the indicated genes were measured by qRT-PCR. Asterisk, *P* < 0.01 compared with Ctrl-si-MEFs. **d**, Impaired induction of IFN- $\beta$  by cytosolic delivery of nucleic acids obtained from various sources at 6 h, but not by stimulation with LPS (200 ng ml<sup>-1</sup>) for 2 h in MEFs deficient in all HMGBs. Asterisk, *P* < 0.01 compared with Ctrl-si-MEFs.

triggers the secretion of IL-1 $\beta$  (ref. 7). The involvement of HMGBs in the DNA- $\text{AIM2}$ -inflammasome pathway was underscored by the observation that B-DNA-induced secretion of IL-1 $\beta$  is significantly impaired in *Hmgb1*<sup>-/-</sup> fetal liver-derived macrophages and RAW264.7 cells expressing the pan-HMGB siRNA (Supplementary Fig. 3a–c). These results suggest that all three HMGBs are required for the full-blown activation of innate immune responses by cytosolic nucleic acids.

We next examined whether the signalling pathways activated through the cytosolic receptors are affected by the absence of HMGBs. The activation of the IRF3 and NF- $\kappa$ B transcription factors was measured in HMGB-si-MEFs or MEFs expressing *Renilla* luciferase-siRNA (Ctrl-si-MEFs) stimulated by B-DNA or poly(I:C).



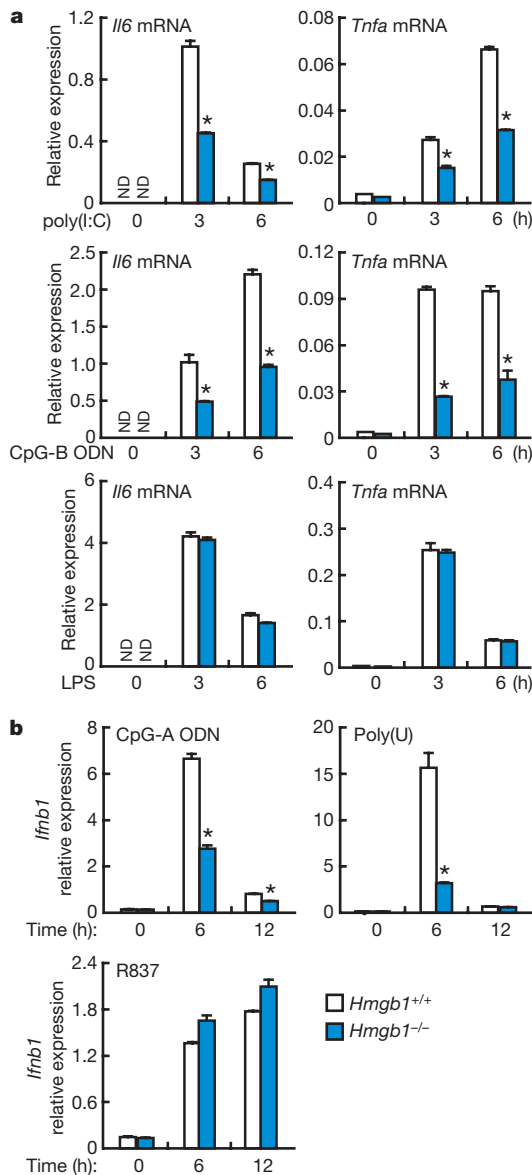
**Figure 2** | A requirement for HMGBs in cytosolic nucleic-acid-receptor-mediated activation of signalling pathways and in antiviral innate immune responses. **a**, **b**, Activation status of IRF3 (**a**) and NF- $\kappa$ B (**b**). MEFs expressing siRNA targeting all HMGBs (HMGB-si) or control siRNA (Ctrl-si) were lipofected with B-DNA or poly(I:C). Dimerization of IRF3 was assessed by native PAGE followed by immunoblot analysis. Activation of NF- $\kappa$ B was analysed by electrophoretic mobility-shift assay. **c**, Induction of type-I IFNs by virus infection. MEFs expressing siRNA targeting all HMGBs (HMGB-si) or control siRNA (Ctrl-si) were infected with VSV or HSV-1. *Ifnb1*, *Ifna1* and *Ifna4* mRNA expression levels were monitored by qRT-PCR. Data in **c** are presented as means and s.d. ( $n = 3$ ). Asterisk,  $P < 0.01$ ; double asterisk,  $P < 0.05$  compared with cells expressing Ctrl-si. ND, not detected.

As shown in Fig. 2a, the formation of dimerized IRF3, a hallmark of IRF3 activation, was strongly suppressed in HMGB-si-MEFs but robust in Ctrl-si-MEFs on stimulation by B-DNA or poly(I:C). Similarly, the activation of extracellular signal-regulated kinase (ERK) (data not shown) and NF- $\kappa$ B was also suppressed in HMGB-si-MEFs (Fig. 2b). Thus, these results indicate that in the context of the nucleic-acid-mediated activation of innate immune responses, HMGBs function in the cytosolic receptor-IRF3/NF- $\kappa$ B signalling pathways.

Because cytosolic nucleic-acid-sensing systems are crucial to anti-viral immunity<sup>1,4</sup>, we next examined the contribution of HMGBs to the virus-induced activation of innate immune responses. The gene induction of type-I IFNs and replication of vesicular stomatitis virus (VSV) or herpes simplex virus type 1 (HSV-1) in infected HMGB-si-MEFs and Ctrl-si-MEFs were measured. *Ifnb1* mRNA induced by these viruses was much lower in HMGB-si-MEFs than in Ctrl-si-MEFs (Fig. 2c). However, the suppression of *Ifna* (*Ifna1* and *Ifna4*) mRNA induction was more resistant to HMGB downregulation than *Ifnb1* mRNA (Fig. 2c), which is presumably due to the operation of one or more mechanisms that do not involve the nucleic acid-HMGB pathway<sup>27</sup>. Accordingly, virus replication was increased in HMGB-si-MEFs (Supplementary Fig. 4a). Similar observations were made for RAW264.7 cells expressing the pan-HMGB siRNA (Supplementary Fig. 4b, c). These data, together with those for virus-derived 'naked' DNAs (Fig. 1d), support the hypothesis that HMGB sensing of virus-derived nucleic acids is critical for the effective avocation of the antiviral innate immune responses.

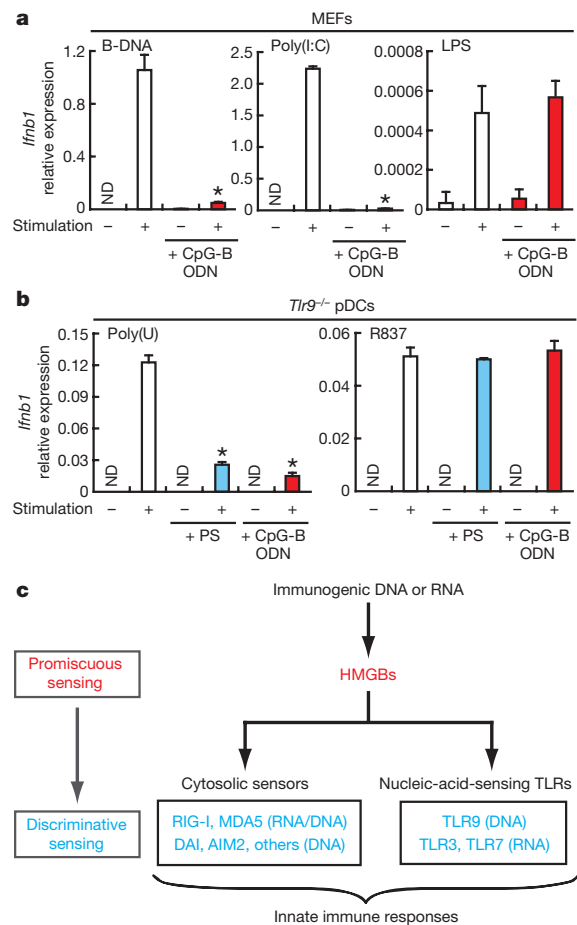
Given that HMGBs (particularly HMGB1) bind to all TLR agonistic nucleic acids (Supplementary Fig. 1b–d), we next examined the roles of HMGBs in the activation of TLRs. A requirement for HMGB1 in TLR9 signalling in cooperation with the multivalent receptor RAGE (receptor for advanced glycation end-products) was reported previously<sup>16,18</sup>. As shown in Fig. 3a, the mRNA induction of proinflammatory cytokines was impaired, albeit partly, in *Hmgb1*<sup>-/-</sup> cDCs on stimulation with TLR3 (poly(I:C)) or TLR9 (CpG-B ODN), whereas mRNA levels were unaffected on stimulation with TLR4 (LPS). The residual induction is likely to be mediated by other HMGBs. Indeed, when the pan-HMGB siRNA was expressed in RAW264.7 cells, mRNA induction by poly(I:C) or CpG-B ODN was more severely inhibited (Supplementary Fig. 5). We also cultured *Hmgb1*<sup>-/-</sup> fetal liver with Flt3-ligand to enrich for plasmacytoid dendritic cells (pDCs), and then stimulated the cells with ligands for TLR7 and TLR9 to drive the pDC-mediated high-level production of type-I IFNs. *Ifnb1* and *Ifna4* mRNA induction by TLR9 agonist CpG-A (D19) ODN<sup>20</sup> was inhibited in *Hmgb1*<sup>-/-</sup> pDCs and CpG-A ODN-stimulated RAW264.7 cells expressing the pan-HMGB siRNA (Fig. 3b and Supplementary Fig. 5). In addition, *Ifnb1* mRNA expression was severely inhibited when *Hmgb1*<sup>-/-</sup> pDCs were stimulated by the TLR7 agonist poly(U), yet when stimulated by R837, which does not bind to HMGBs (Supplementary Fig. 1b), the expression levels were the same as those in wild-type pDCs (Fig. 3b). The deficiency in HMGB1 had a more profound effect on TLR7 stimulation by poly(U) than on TLR9 stimulation by CpG-A ODN (Fig. 3b), possibly because poly(U) is not bound by HMGB2 (Supplementary Fig. 1b). We speculate that, like TLR9 signalling, TLR3 and TLR7 signalling may also involve RAGE-bound HMGB1 (ref. 18). Thus, these results together underscore the role of HMGBs in the activation of all nucleic-acid-sensing TLRs.

Because CpG-B ODN and the antagonist PS bind to HMGBs with extraordinarily high affinities (Supplementary Fig. 1b, c), we tested whether these compounds interfere with innate immune responses evoked by immunogenic nucleic acids of lower affinities; we reasoned that a functional consequence of the pretreatment of HMGBs with these high-binding-affinity nucleotide analogues would 'mask' HMGBs, thereby selectively interfering with the activation of innate immune responses by other nucleic acids. Indeed, wild-type MEFs were hyporesponsive to stimulation with cytosolic B-DNA or poly(I:C) when pretreated with CpG-B ODN, which itself does not evoke a detectable



**Figure 3 | A requirement for HMGB1 in the nucleic-acid-mediated activation of TLRs.** *Hmgb1*<sup>+/+</sup> or *Hmgb1*<sup>-/-</sup> cDCs (**a**) and pDCs (**b**) were stimulated by the indicated TLR ligands, and then mRNA expression levels for various cytokines were measured by qRT-PCR. Data in all panels are presented as means and s.d. ( $n = 3$ ). Asterisk,  $P < 0.01$  compared with wild-type cells. ND, not detected.

cytokine response in this cell type, whereas their response to LPS remained normal (Fig. 4a). As expected, the suppressive effect of CpG-B ODN was weaker if added after stimulation with B-DNA (Supplementary Fig. 6a). Furthermore, even in TLR9-deficient pDCs and cDCs, a profound hyporesponse was observed on pretreatment with CpG-B ODN or PS followed by stimulation by TLR3 or TLR7 with their cognate nucleic acid ligands (Fig. 4b). In contrast, activation of TLR7 by R837 was not affected by pretreatment with CpG-B ODN or PS (Fig. 4b and Supplementary Fig. 6b). Consistent with this, when B-DNA was used as an adjuvant the induction of ovalbumin-specific CD8 T-cell responses was severely impaired by the pretreatment of TLR9-deficient mice with CpG-B (M. Miyajima and T.T., unpublished observation). These results therefore lend further support to the universal role of HMGBs in the nucleic-acid-mediated innate immune responses, although strictly the possibility is not ruled out that the CpG-B ODN-mediated or PS-mediated interference with nucleic acid signalling described here may involve additional mechanisms.



**Figure 4 | Interference of HMGB-nucleic-acid-induced activation of innate immune responses by use of high-binding-affinity nucleotide analogues.**

**a**, MEFs were stimulated by cytosolic delivery of B-DNA, poly(I:C) or LPS with or without pretreatment for 30 min with 1  $\mu$ M CpG-B ODN. *Ifnb1* mRNA expression levels were then monitored by qRT-PCR. Data in all panels are means and s.d. ( $n = 3$ ). Asterisk,  $P < 0.01$ , pretreated cells versus non-pretreated cells. ND, not detected. **b**, Bone marrow-derived *Tlr9*<sup>-/-</sup> pDCs were either lipofected with 1  $\mu$ g ml<sup>-1</sup> poly(U) or stimulated with 25  $\mu$ g ml<sup>-1</sup> R837 for 8 h, in the absence of pretreatment or after pretreatment with 5  $\mu$ M PS or 1  $\mu$ M CpG-B ODN for 30 min. The mRNA expression levels of the indicated genes were measured by qRT-PCR. Asterisk,  $P < 0.01$ , pretreated cells versus non-pretreated cells. **c**, A schematic view of the hierarchy in the nucleic-acid-mediated activation of immune responses. All immunogenic nucleic acids bind HMGBs (promiscuous sensing), which is required for subsequent recognition by specific pattern recognition receptors (discriminative sensing) to activate the innate immune responses.

We propose, and our data support, that HMGBs serve as universal sentinels for nucleic acids that are required for the full-blown, nucleic-acid-induced activation of innate immune responses mediated by the more discriminative pattern recognition receptors (Fig. 4c). What is the biological significance of the HMGB's function that we describe here? Because HMGBs are broadly found in multicellular organisms including plants<sup>28</sup>, one may speculate that HMGBs are a remnant of a primitive sensing system for foreign nucleic acids, which has evolved along with and in support of discriminative TLRs and cytosolic receptors of the innate immune signalling system. In this regard, it is worth noting that HMGB1 associates with TLR9 (refs 16, 18) and RLRs (Z.W. and T.T., unpublished observation), suggesting that nucleic-acid-bound HMGBs have evolved to function as co-ligands for discriminative nucleic-acid receptors; in other words, the binding of nucleic acids to HMGBs is a precondition for the more efficient, subsequent recognition by and activation of TLR, RLR and cytosolic nucleic-acid receptors (Fig. 4c).



Nevertheless, it remains unclear exactly where HMGBs bind to nucleic acids and how the HMGB–nucleic acid complexes activate their respective receptor signalling cascades. *In vivo* visualization data lend support to the concept that HMGB1 functions together with RIG-I on the endosomal membrane, through which nucleic acids are released into the cytosol by low-pH-dependent membrane fusion<sup>29,30</sup> (Supplementary Fig. 7). In view of a previous report showing the interaction between HMGB1 and TLR9 in the endosomal compartment<sup>16</sup>, we conjecture that endosome membrane-associated signalling may be universal to nucleic-acid-mediated innate immune responses. In addition to their role in nucleic-acid sensing described here, HMGBs serve additional functions including gene transcription, all of which may be viewed as a long-term evolutionary consequence of their ability to bind nucleic acids. Nevertheless, further study is required for a more complete understanding of the precise role and mechanism of action of these proteins. Finally, our findings have a number of interesting therapeutic implications, such as in the use of HMGB inhibitory compounds in the suppression of nucleic-acid-mediated pathological immune responses.

## METHODS SUMMARY

**Pull-down assay.** MEFs were treated with poly(dA:dT)•(dT-dA) (B-DNA; 10 µg ml<sup>-1</sup>) for 4 h before mass spectrometry analysis. After stimulation, cells were homogenized with a Dounce homogenizer (Wheaton Science) in homogenization buffer (20 mM HEPES pH 7.4, 20% glycerol, 50 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM phenylmethylsulphonyl fluoride (PMSF), 10 µg ml<sup>-1</sup> aprotinin, 10 µg ml<sup>-1</sup> leupeptin). Cytosolic protein extract was prepared from the homogenate by centrifugation at 19,000g for 30 min, and then incubated with 1.4 µg ml<sup>-1</sup> B-DNA, the 5' terminus of which had been conjugated with biotin, and streptavidin (SA)-conjugated magnetic beads (Invitrogen) at 4 °C for 15 min. The 'pulled-down' sample was treated with DNase I (Invitrogen) in reaction buffer (20 mM Tris-HCl pH 8.4, 20 mM MgCl<sub>2</sub>, 50 mM KCl) and the resultant supernatant was subjected to silver staining (Invitrogen) or mass spectrometry analysis. The *in vitro* pull-down assay was performed as essentially described previously<sup>8</sup>. Recombinant HMGB1, HMGB2 and HMGB3 proteins were first incubated for 30 min at 15–20 °C with or without competitors. The supernatants were then mixed for 30 min at 4 °C with biotin-conjugated B-DNA after preincubation with SA-conjugated magnetic beads in binding buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Nonidet P40, 100 µg ml<sup>-1</sup> leupeptin, 1 mM PMSF, 1 mM Na<sub>3</sub>VO<sub>4</sub>). The mixture was then washed extensively with the same buffer, separated by SDS-PAGE and immunoblotted with anti-HMGB1, anti-HMGB2 and anti-HMGB3 antibodies.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 26 August; accepted 17 September 2009.

- Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. *Cell* **124**, 783–801 (2006).
- Medzhitov, R. Recognition of microorganisms and activation of the immune response. *Nature* **449**, 819–826 (2007).
- Hornung, V. *et al.* Sequence-specific potent induction of IFN- $\alpha$  by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nature Med.* **11**, 263–270 (2005).
- Kato, H. *et al.* Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* **441**, 101–105 (2006).
- Yoneyama, M. *et al.* The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature Immunol.* **5**, 730–737 (2004).
- Ishii, K. J. *et al.* TANK-binding kinase-1 delineates innate and adaptive immune responses to DNA vaccines. *Nature* **451**, 725–729 (2008).
- Schroder, K., Muruve, D. A. & Tschopp, J. Innate immunity: cytoplasmic DNA sensing by the AIM2 inflammasome. *Curr. Biol.* **19**, R262–R265 (2009).
- Takaoka, A. *et al.* DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature* **448**, 501–505 (2007).
- Choi, M. K. *et al.* A selective contribution of the RIG-I-like receptor pathway to type I interferon responses activated by cytosolic DNA. *Proc. Natl Acad. Sci. USA* advance online publication, doi:10.1073/pnas.0910983106 (1 October 2009).
- Chiu, Y. H., Macmillan, J. B. & Chen, Z. J. RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. *Cell* **138**, 576–591 (2009).

- Ablasser, A. *et al.* RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate. *Nature Immunol.* **10**, 1065–1072 (2009).
- Ishii, K. J. *et al.* A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. *Nature Immunol.* **7**, 40–48 (2006).
- Bianchi, M. E. & Manfredi, A. A. High-mobility group box 1 (HMGB1) protein at the crossroads between innate and adaptive immunity. *Immunol. Rev.* **220**, 35–46 (2007).
- Scaffidi, P., Misteli, T. & Bianchi, M. E. Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. *Nature* **418**, 191–195 (2002).
- Apetoh, L. *et al.* The interaction between HMGB1 and TLR4 dictates the outcome of anticancer chemotherapy and radiotherapy. *Immunol. Rev.* **220**, 47–59 (2007).
- Ivanov, S. *et al.* A novel role for HMGB1 in TLR9-mediated inflammatory responses to CpG-DNA. *Blood* **110**, 1970–1981 (2007).
- Kazama, H. *et al.* Induction of immunological tolerance by apoptotic cells requires caspase-dependent oxidation of high-mobility group box-1 protein. *Immunity* **29**, 21–32 (2008).
- Tian, J. *et al.* Toll-like receptor 9-dependent activation by DNA-containing immune complexes is mediated by HMGB1 and RAGE. *Nature Immunol.* **8**, 487–496 (2007).
- Krieg, A. M. *et al.* CpG motifs in bacterial DNA trigger direct B-cell activation. *Nature* **374**, 546–549 (1995).
- Klinman, D. M. Immunotherapeutic uses of CpG oligodeoxynucleotides. *Nature Rev. Immunol.* **4**, 249–258 (2004).
- Haas, T. *et al.* The DNA sugar backbone 2' deoxyribose determines Toll-like receptor 9 activation. *Immunity* **28**, 315–323 (2008).
- Nemeth, M. J. *et al.* Hmgb3: an HMG-box family member expressed in primitive hematopoietic cells that inhibits myeloid and B-cell differentiation. *Blood* **102**, 1298–1306 (2003).
- Vaccari, T., Beltrame, M., Ferrari, S. & Bianchi, M. E. *Hmg4*, a new member of the *Hmg1/2* gene family. *Genomics* **49**, 247–252 (1998).
- Stetson, D. B. & Medzhitov, R. Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. *Immunity* **24**, 93–103 (2006).
- Pichlmair, A. & Reis e Sousa, C. Innate recognition of viruses. *Immunity* **27**, 370–383 (2007).
- Hornung, V. *et al.* 5'-Triphosphate RNA is the ligand for RIG-I. *Science* **314**, 994–997 (2006).
- Paladino, P., Cummings, D. T., Noyce, R. S. & Mossman, K. L. The IFN-independent response to virus particle entry provides a first line of antiviral defense that is independent of TLRs and retinoic acid-inducible gene I. *J. Immunol.* **177**, 8008–8016 (2006).
- Grasser, K. D., Launholt, D. & Grasser, M. High mobility group proteins of the plant HMGB family: dynamic chromatin modulators. *Biochim. Biophys. Acta* **1769**, 346–357 (2007).
- Kampmann, T., Mueller, D. S., Mark, A. E., Young, P. R. & Kobe, B. The role of histidine residues in low-pH-mediated viral membrane fusion. *Structure* **14**, 1481–1487 (2006).
- Weissenhorn, W., Hinz, A. & Gaudin, Y. Virus membrane fusion. *FEBS Lett.* **581**, 2150–2155 (2007).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank C. Reis e Sousa and J. Rehwinkel for 5'-triphosphate RNA; Y. Kawaguchi for HSV-1 DNA; A. Katoh and M. Kidokoro for the vaccinia virus (MO) genome for ligand stimulation; J. Vilcek, R. Medzhitov, R. Schekman, J. V. Ravetch and H. Ohno for invaluable advice; and R. Takeda, M. Shishido, Y. Nakaima and T. Oh for technical assistance. This work was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas 'Integrative Research Toward the Conquest of Cancer,' Grant-in-Aid for Scientific Research (A), and Global Center of Excellence Program 'Integrative Life Science Based on the Study of Biosignaling Mechanisms' from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and by the Korea Science and Engineering Foundation Grant. T.B. and Z.W. are research fellows of the Japan Society for the Promotion of Science.

**Author Contributions** H.Y., T.B., Z.W. and M.C. designed, performed and analysed most of the experiments in this study. T. Kawamura and T. Kodama collected data and performed the mass spectrometry analysis. H.N. performed experiments on RAW264.7 cells. M.N. performed the fluorescence microscopy analysis. Y.L. and R.K. cloned the constructs. S.H. assisted with the virus infection experiments. A.S., L.R. and M.B. provided precious materials and invaluable advice. K.H., T. Tamura and T. Taniguchi provided overall coordination with respect to conception, design and supervision of the study. H.Y., T.B., D.S., T. Tamura and T. Taniguchi wrote the manuscript with comments from the co-authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to T. Taniguchi ([tada@m.u-tokyo.ac.jp](mailto:tada@m.u-tokyo.ac.jp)).

## METHODS

**Mice, cells and reagents.** Mice used for this study were on the genetic background of C57BL/6, except *Tlr9*<sup>-/-</sup> mice, which were on the Balb/c background. The generation of *Tlr9*<sup>-/-</sup>, *Hmgb1*<sup>-/-</sup> and *Hmgb2*<sup>-/-</sup> mice were described previously<sup>12,31,32</sup>. MEFs, RAW264.7, NIH/3T3 and HEK293T cells, bone marrow-derived cDCs and pDCs from *Tlr9*<sup>-/-</sup> mice were maintained as described previously<sup>8,33</sup>. *Hmgb1*<sup>-/-</sup> macrophages, cDCs and pDCs were generated from fetal liver haematopoietic progenitors (lineage-marker-negative cells purified with the MACS Lineage depletion kit from Miltenyi Biotec) by culturing fetal liver cells for 2 days in the presence of SCF (20 ng ml<sup>-1</sup>), IL-3 (10 ng ml<sup>-1</sup>) and IL-6 (10 ng ml<sup>-1</sup>) followed by a 6-day culture in the presence of 20 ng ml<sup>-1</sup> macrophage colony-stimulating factor (for macrophages), 20 ng ml<sup>-1</sup> GM-CSF (for cDCs) or 100 ng ml<sup>-1</sup> human Flt3-ligand (for pDCs). SCF, IL-3 and IL-6 were purchased from Peprotech. IFN- $\gamma$  and TNF- $\alpha$  were purchased from R&D Systems. IFN- $\beta$  was provided by TORAY Industries. B-DNA and calf thymus genomic DNA were purchased from Sigma. Biotin-conjugated poly(dA-dT) $\cdot$ poly(dT-dA) was purchased from Hokkaido System Science. Other oligo DNAs including ISD<sup>24</sup>, CpG ODNs, fluorescein isothiocyanate-conjugated base-free phosphorothioate deoxyribose homopolymer (PS; 20-mer) and fluorescein isothiocyanate-conjugated base-free natural deoxyribose homopolymer (PD; 20-mer) were purchased from Fasmac. *E. coli* DNA and R837 were purchased from InvivoGen. Poly(U) and LPS were purchased from Sigma. Poly(I:C) was purchased from GE Healthcare Biosciences. B-DNA, poly(I:C) and other nucleic acid ligands were used at a concentration of 10  $\mu$ g ml<sup>-1</sup> as described previously<sup>8</sup>, unless otherwise mentioned. The complex formation of CpG-A ODN with DOTAP (Roche) was performed as described previously<sup>33</sup>. MitoTracker Deep Red 633 was purchased from Invitrogen. Antibodies against the following proteins were used: HMGB1 and HMGB2 (Abcam), HMGB3 (TransGenic Inc.), IRF3 (ZM3; Zymed),  $\beta$ -actin (AC-15; Sigma-Aldrich), NF- $\kappa$ B p65 (C20; Santa Cruz Biotechnology) and phospho-STAT1 (58D6; Cell Signaling).

**Plasmid constructions and protein purification.** Murine HMGB cDNAs were obtained by RT-PCR on total RNA from MEFs, and then cloned into pGEX4T3 vector (GE Healthcare Biosciences) at the *Sal*I and *Not*I sites. The glutathione S-transferase (GST)-tagged HMGB proteins were purified by using glutathione-Sepharose beads (GE Healthcare Biosciences). HMGB proteins and GST protein were separated by treatment with thrombin proteinase (Novagen). Other expression vectors are listed in the Supplementary Information.

**Immunoblot analysis.** Cell lysis and immunoblot analysis were performed as described previously<sup>8</sup>. IRF3 dimer was assessed by native PAGE, followed by immunoblot analysis with anti-mouse IRF3 antibody as described previously<sup>8</sup>.

The quantification of IRF3 dimer was performed by the NIH Image application. Similar results were obtained in three independent experiments.

**RNA analysis.** RNA extraction and reverse transcription were performed as described previously<sup>8</sup>. qRT-PCR analysis was performed with a Lightcycler480 and SYBR Green system (Roche Molecular Biochemicals). All data are presented as relative expression units after normalization to GAPDH. Data are presented as means and s.d. for triplicate determinations. All data were reproduced in two or more additional independent experiments. Primer sequences for murine GAPDH, IL-6, RANTES, I $\kappa$ B- $\alpha$ , IFN- $\alpha$ 1, IFN- $\alpha$ 4 and IFN- $\beta$  have been described previously<sup>8</sup>. Other primer sequences are listed in the Supplementary Information.

**Statistical analysis.** Differences between control and experimental groups were evaluated with Student's *t*-test.

**Enzyme-linked immunosorbent assay (ELISA).** Murine IFN- $\beta$ , IL-6 or IL-1 $\beta$  was measured by ELISA. IFN- $\beta$  ELISA kit was purchased from PBL Biomedical Laboratories. IL-6 and IL-1 $\beta$  ELISA kits were obtained from R&D Systems. All data were reproduced in two additional independent experiments.

**RNA interference.** siRNA vectors were constructed by inserting oligonucleotides into *Eco*RI and *Hind*III sites of the pSUPER.retro.puro retrovirus vector<sup>9</sup>. The siRNA targeting sequences for murine HMGB1/2/3 (pan-HMGB siRNA), HMGB2 and *Renilla* luciferase (control) are 5'-GTATGAGAAGGATATTGCT-3', 5'-GCGTTACGAGAAACCAGTT-3' and 5'-GTAGCGCGGTGTATTATACA-3', respectively. Retroviruses were prepared as described previously<sup>9</sup>. Transduced cells were selected by puromycin (Sigma) at 2  $\mu$ g ml<sup>-1</sup> (MEFs) or 4  $\mu$ g ml<sup>-1</sup> (RAW264.7 cells) for 48 h.

**Electrophoretic mobility-shift assay.** Electrophoretic mobility-shift assay was performed as described previously<sup>8</sup>. An oligonucleotide probe containing a consensus NF- $\kappa$ B binding sequence was used<sup>8</sup>. The presence of p65 in the NF- $\kappa$ B-DNA binding complex was also confirmed by detection of a supershifted band with an anti-p65 antibody (data not shown).

**Viral infection.** Cells were infected for 12 h with 1.0 multiplicity of infection of HSV-1 or VSV, as described previously<sup>8</sup>. To measure the yield of HSV-1 or VSV, a plaque-forming assay was performed as described previously<sup>8</sup>. All data were reproduced in two additional independent experiments. Virus preparation was described previously<sup>8</sup>.

31. Calogero, S. *et al.* The lack of chromosomal protein Hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice. *Nature Genet.* **22**, 276–280 (1999).
32. Ronfani, L. *et al.* Reduced fertility and spermatogenesis defects in mice lacking chromosomal protein Hmgb2. *Development* **128**, 1265–1273 (2001).
33. Honda, K. *et al.* Spatiotemporal regulation of MyD88-IRF-7 signalling for robust type-I interferon induction. *Nature* **434**, 1035–1040 (2005).

## LETTERS

Requirement for NF- $\kappa$ B signalling in a mouse model of lung adenocarcinomaEtienne Meylan<sup>1</sup>, Alison L. Dooley<sup>1</sup>, David M. Feldser<sup>1</sup>, Lynn Shen<sup>1</sup>, Erin Turk<sup>1</sup>, Chensi Ouyang<sup>1</sup> & Tyler Jacks<sup>1</sup>

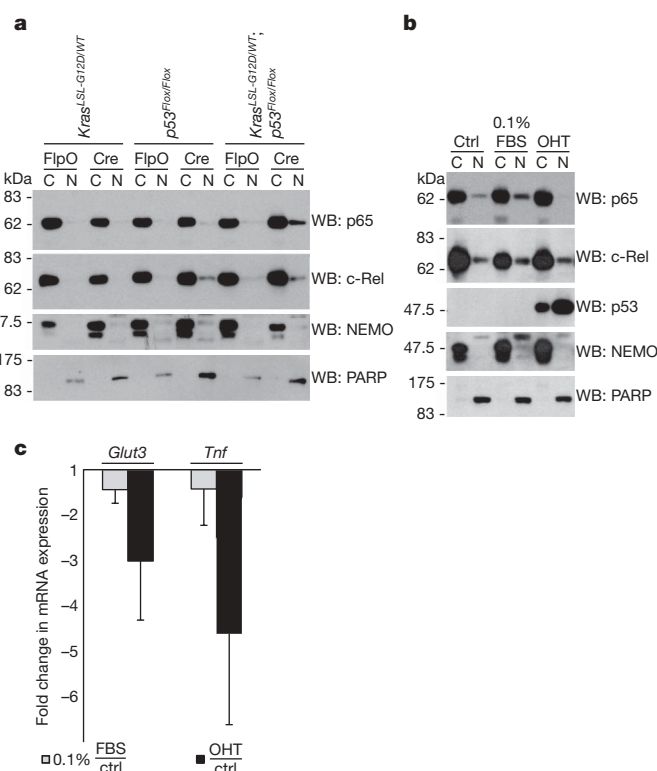
NF- $\kappa$ B transcription factors function as crucial regulators of inflammatory and immune responses as well as of cell survival<sup>1</sup>. They have also been implicated in cellular transformation and tumorigenesis<sup>2–6</sup>. However, despite extensive biochemical characterization of NF- $\kappa$ B signalling during the past twenty years, the requirement for NF- $\kappa$ B in tumour development *in vivo*, particularly in solid tumours, is not completely understood. Here we show that the NF- $\kappa$ B pathway is required for the development of tumours in a mouse model of lung adenocarcinoma. Concomitant loss of p53 (also known as Trp53) and expression of oncogenic Kras(G12D) resulted in NF- $\kappa$ B activation in primary mouse embryonic fibroblasts. Conversely, in lung tumour cell lines expressing Kras(G12D) and lacking p53, p53 restoration led to NF- $\kappa$ B inhibition. Furthermore, the inhibition of NF- $\kappa$ B signalling induced apoptosis in p53-null lung cancer cell lines. Inhibition of the pathway in lung tumours *in vivo*, from the time of tumour initiation or after tumour progression, resulted in significantly reduced tumour development. Together, these results indicate a critical function for NF- $\kappa$ B signalling in lung tumour development and, further, that this requirement depends on p53 status. These findings also provide support for the development of NF- $\kappa$ B inhibitory drugs as targeted therapies for the treatment of patients with defined mutations in Kras and p53.

More than one million people worldwide die from lung cancer each year, making it the leading cause of cancer mortality in both women and men. Non-small-cell lung cancer (NSCLC) accounts for 80% of all lung cancer cases, with adenocarcinoma being the major subtype. NSCLC development is associated with frequent mutations in a few well-defined oncogenes and tumour suppressor genes. Oncogenic Kras mutations occur in approximately 20–30% of patients with lung adenocarcinoma, whereas inactivating mutations in the tumour suppressor protein p53 are found in at least 50% of NSCLC cases<sup>7</sup>.

The overexpression of oncogenic forms of Ras has been reported to result in NF- $\kappa$ B activation<sup>3,8</sup>, whereas wild-type p53 has been shown to antagonize NF- $\kappa$ B activity<sup>9–14</sup>. Therefore, we proposed that induction of endogenous Kras(G12D) or p53 deletion might affect the NF- $\kappa$ B signalling pathway. We used primary mouse embryonic fibroblasts (MEFs) carrying either the Cre-activatable Kras(G12D) allele (*Kras*<sup>LSL-G12D/WT</sup>), in which LSL denotes *Lox-stop-Lox*, a conditional loss-of-function p53 allele (*p53*<sup>Flox/Flox</sup>), or the two in combination (*Kras*<sup>LSL-G12D/WT</sup>; *p53*<sup>Flox/Flox</sup>)<sup>15–17</sup>. Cells were infected with adenoviruses expressing either Cre (to activate oncogenic Kras and/or inactivate p53) or FlpO recombinase (as a control). As shown in Fig. 1a, activation of Kras(G12D) with concomitant loss of p53 led to an increased nuclear accumulation of the NF- $\kappa$ B transcription factor subunit p65 (also known as RelA), a prerequisite for NF- $\kappa$ B activity. In contrast, p65 remained exclusively cytoplasmic after Kras(G12D) expression in wild-type p53-expressing cells or after loss

of p53 in the absence of Kras(G12D) expression, suggesting that the combination of mutations is required for enhanced NF- $\kappa$ B activation, at least in these cells. Of note, another NF- $\kappa$ B subunit, c-Rel, showed a slight increase in nuclear localization after p53 loss, both alone and after Kras(G12D) expression (Fig. 1a), perhaps reflecting crosstalk between p53 and c-Rel.

We next addressed whether restoration of p53 expression in p53-mutant cells might influence p65 nuclear localization. We used Kras(G12D)-expressing lung adenocarcinoma cell lines with a Cre-activatable wild-type p53 allele, which were derived from lung



**Figure 1 | NF- $\kappa$ B activity depends on p53.** **a**, MEFs were infected with adenoviruses expressing Cre or FlpO recombinase. After 6 days, cytoplasmic (C) and nuclear (N) lysates were analysed by western blot (WB). NEMO (cytoplasmic) and PARP (nuclear) were used to determine purity. **b**, **c**, *Kras*<sup>LA2/WT</sup>; *p53*<sup>LSL/LSL</sup>; *Cre-ER*<sup>T2</sup>-lung-tumour-derived cells were untreated (ctrl), cultured in 0.1% FBS or stimulated with 200 nM 4-hydroxytamoxifen (OHT). After 4 days, cells were lysed and analysed as in **a** (**b**), or RNA was purified, reverse transcribed, and complementary DNA was amplified using probes for *Glut3*, *Tnf* or *Gapdh* (internal control) (**c**). Results are means and s.d. ( $n = 3$  independent cell lines), presented as fold changes (treatment divided by control; see Methods)<sup>30</sup>.

<sup>1</sup>Koch Institute for Integrative Cancer Research, and Department of Biology, and Howard Hughes Medical Institute, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.



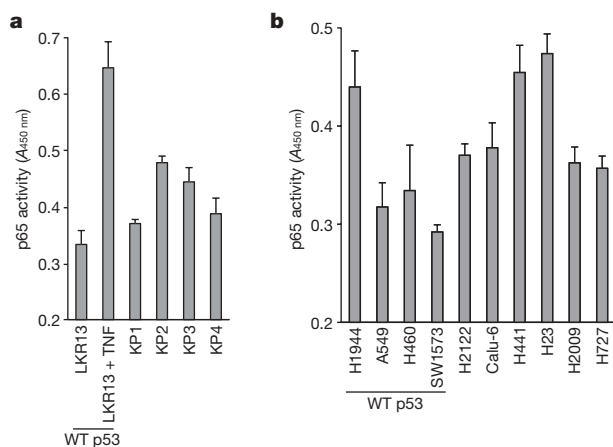
tumours of mice generated by the crossing of *Kras*<sup>LA2/WT</sup> (ref. 18) and *p53*<sup>LSL/LSL</sup>; *Cre-ER*<sup>T2</sup> (ref. 19) mice (D.M.F. and T.J., unpublished data). In the absence of p53 restoration, nuclear p65 was evident in these cells, consistent with the data from MEFs (see Fig. 1a). After 4-hydroxytamoxifen-induced p53 restoration, p65 nuclear localization was diminished (Fig. 1b). In contrast, the nuclear levels of c-Rel remained unchanged. Reduced levels of nuclear p65 were not simply a consequence of cell cycle arrest, because p65 remained detectable in the nucleus of cells arrested by culturing in 0.1% serum. Notably, restoration of p53 and decreased nuclear p65 expression were associated with reduced expression of the NF- $\kappa$ B-target gene *Tnf* (Fig. 1c). The expression of *Glut3*, which contributes to and is upregulated by NF- $\kappa$ B activity in p53-deficient MEFs<sup>10</sup>, was also reduced after restoration of p53 (Fig. 1c).

Using enzyme-linked immunosorbent assay (ELISA), we directly monitored NF- $\kappa$ B activity in lung adenocarcinoma cell lines derived from tumours generated in the conditional compound knock-in (*Kras*<sup>LSL-G12D/WT</sup>; *p53*<sup>Flox/Flox</sup>) mouse. These mice (hereafter referred to as KP mice) develop advanced lung cancers that closely recapitulate the development of human lung adenocarcinoma<sup>15</sup>. In KP tumour cell lines, NF- $\kappa$ B p65 DNA-binding activity was consistently higher than in cell lines derived from the *Kras*<sup>LA2/WT</sup> allele (LKR), which also express *Kras*(G12D) but retain wild-type p53 expression (Fig. 2a). We also examined NF- $\kappa$ B activity in human NSCLC cell lines that contain gain-of-function mutations in *Kras* but differ in their p53 status. Consistent with the mouse data, NF- $\kappa$ B p65 DNA-binding activity was increased in p53 mutant compared to wild-type p53 cell lines, with one exception (Fig. 2b). Notably, nuclear p65 was higher in the p53 mutant cells (Supplementary Fig. 1). Hence, canonical NF- $\kappa$ B signalling is increased in lung tumour-derived cells with altered p53.

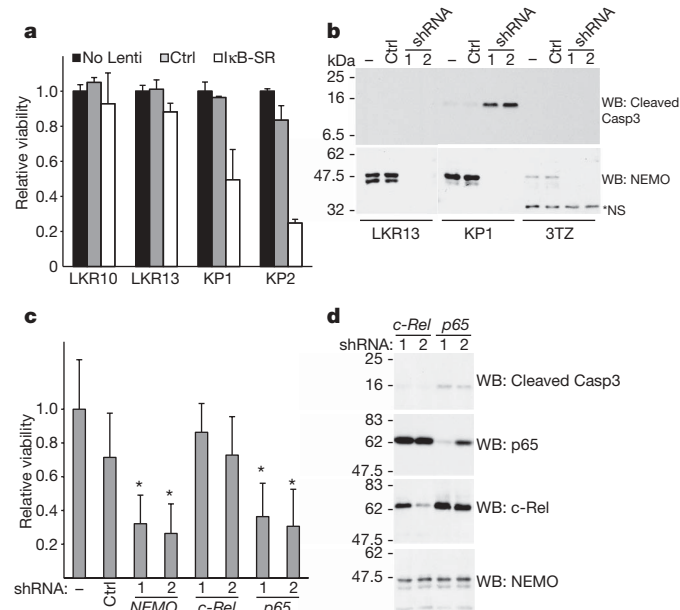
To determine whether NF- $\kappa$ B provides pro-survival signals in KP cell lines, we examined the consequences of pathway inhibition. Cell viability was monitored after infection of cells with lentiviruses expressing a non-phosphorylatable, dominant-negative form of I $\kappa$ B $\alpha$  (encoded by *Nfkbia*; super repressor, I $\kappa$ B-SR). As shown in Fig. 3a, the expression of I $\kappa$ B-SR led to reduced KP cell viability. In contrast, the two LKR cell lines expressing mutant *Kras* and wild-type p53 (LKR10 and LKR13) were insensitive to I $\kappa$ B-SR expression. These results are consistent with the data showing increased NF- $\kappa$ B activity in KP cells (see Fig. 2a). NF- $\kappa$ B signalling was also inhibited by short hairpin RNA (shRNA)-mediated knockdown of *NEMO* (also known as *Ikbkg*), an essential factor in the canonical NF- $\kappa$ B pathway<sup>20,21</sup>. Infection of LKR, KP and 3TZ fibroblast cell lines with pSicoRev retroviral vectors<sup>22</sup> expressing one of two hairpins targeting *NEMO* resulted in efficient knockdown in all cases, whereas apoptosis

(as measured by cleaved caspase-3 detection) was only triggered in the KP cell line (Fig. 3b). To establish whether one of the two NF- $\kappa$ B subunits, p65 or c-Rel, transmitted these anti-apoptotic effects, we used shRNA-mediated knockdown to inhibit them in the same cells. Knockdown of p65, but not of c-Rel, led to reduced cell viability and cleavage of caspase-3 (Fig. 3c, d). These results demonstrate that in the context of oncogenic *Kras* expression and p53 mutation, the canonical NEMO-, I $\kappa$ B- and p65-dependent NF- $\kappa$ B signalling pathway is required for survival of lung cell lines.

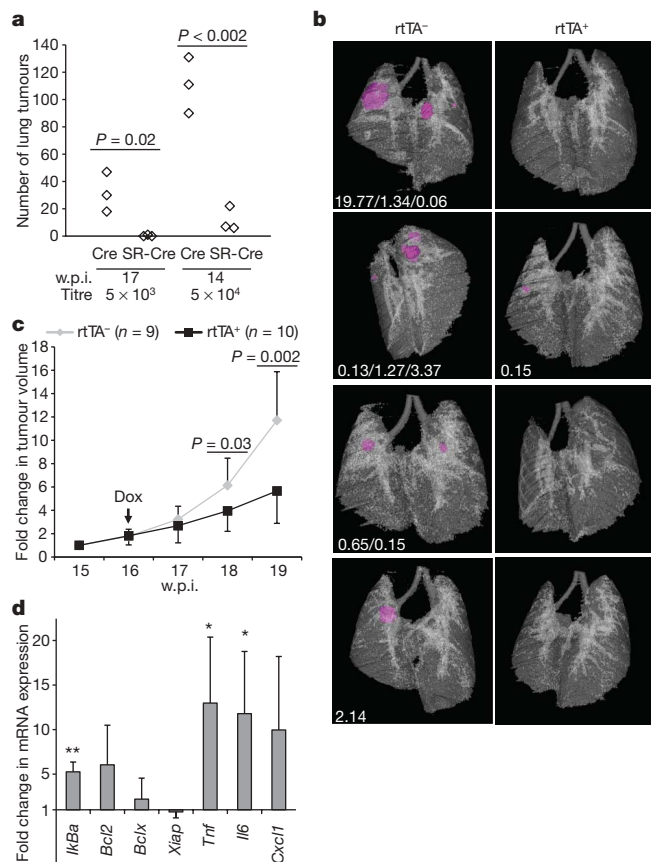
To understand the relevance of these findings *in vivo*, we examined the NF- $\kappa$ B signalling pathway during lung adenocarcinoma development in KP mice. Tumours from KP mice had enhanced NF- $\kappa$ B p65 nuclear localization by immunohistochemistry when compared to tumours from *Kras*(G12D) mice (*Kras*<sup>LSL-G12D/WT</sup>; Supplementary Fig. 2). To test the functional requirement for NF- $\kappa$ B in tumour initiation and development, we infected KP mice with a bifunctional lentivirus capable of expressing both I $\kappa$ B-SR and Cre constitutively from independent promoters or a control lentivirus carrying only Cre (Supplementary Fig. 3a). Each mouse was either infected with low- ( $5 \times 10^3$ ) or high- ( $5 \times 10^4$ ) titre lentiviral particles. Seventeen weeks after infection with low-titre control viruses, an average of 30 tumours was detected per mouse (Fig. 4a, left), whereas mice infected with lentiviruses expressing Cre and I $\kappa$ B-SR had either a single tumour or none at all. Using the high-titre lentiviruses, a similarly notable difference between Cre- and I $\kappa$ B-SR-Cre-infected mice was observed (90–131 tumours compared to 6–22; Fig. 4a, right). Thus, the inhibition of NF- $\kappa$ B at the time of tumour initiation in the KP mouse model had a profound effect on subsequent tumour development. To determine whether NF- $\kappa$ B was required to maintain the growth and/or viability of established tumours, we cloned I $\kappa$ B-SR into a lentiviral construct downstream of a tetracycline-response element (TRE), to generate a



**Figure 2 | NF- $\kappa$ B p65 DNA-binding activity is increased in NSCLC cells with altered p53.** **a**, **b**, Five micrograms of nuclear lysates of the indicated mouse (**a**) or human (**b**) cell lines were analysed for p65 DNA-binding activity by an ELISA (see Methods). Cells containing wild-type p53 are indicated. As a positive control of p65 activity, LKR13 cells (**a**) were stimulated with TNF for 30 min before lysis. Data in **a** and **b** are means and s.d. ( $n = 3$ ).



**Figure 3 | NF- $\kappa$ B inhibition results in apoptosis of p53-deficient lung adenocarcinoma cells.** **a**, Two wild-type p53 (LKR10 and LKR13) and two *Kras*<sup>LSL-G12D/WT</sup>; *p53*<sup>Flox/Flox</sup> (KP) mouse cell lines were uninfected (no Lenti), or infected with control or I $\kappa$ B-SR-expressing lentiviruses. After 72 h, cells were assayed for viability (see Methods). Data are means and s.d. ( $n = 3$ ) of cell viability relative to uninfected cells (set to 1). **b**, Lung cell lines or fibroblasts (3TZ) were untreated (–) or infected with pSicoRev expressing a *NEMO* (shRNA1/2) or control hairpin. After 72 h, cell extracts were analysed by western blot. \*NS denotes non-specific staining. **c**, **d**, KP cells were infected with pSicoRev expressing the indicated hairpin, or were uninfected (–). After 72 h, cells were assayed for viability (**c**), or cell extracts were analysed by western blot (**d**). Data in **c** are means and s.d. ( $n = 6$ ) of cell viability relative to uninfected cells (set to 1). \* $P < 0.05$  relative to control hairpin.



**Figure 4 | NF-κB inhibition impairs lung adenocarcinoma development.**

**a**,  $Kras^{LSL-G12D/WT}$ ;  $p53^{Flx/Flx}$  mice were infected with Lenti-Pgk.Cre (Cre) or Lenti-UbC.IκB-SR; Pgk.Cre (SR-Cre) viruses, euthanized 17 (left) or 14 (right) weeks post-infection (w.p.i.) and the tumour number was counted (see Methods). **b–d**,  $Kras^{LSL-G12D/WT}$ ;  $p53^{Flx/Flx}$ ; CCSP-rtTA<sup>+</sup> ( $rtTA^{+/+}$ ) or -rtTA<sup>-</sup> ( $rtTA^{-/-}$ ) mice were infected with Lenti-TRE.IκB-SR; Pgk.Cre viruses ( $10^4$  viral particles). **b**, From the day of infection, mice were fed a Dox diet. At 15 w.p.i., μ-CT imaged tumours (magenta) and lungs were reconstructed and individual tumour volumes (mm<sup>3</sup>, white text) were measured. **c**, From 15–19 w.p.i., lungs were μ-CT imaged, and individual tumour volumes were measured. Data points represent means of fold change and s.d. in tumour volume relative to 15 w.p.i. (set to 1). **d**, After 1 week of Dox treatment, RNA was isolated from tumours of  $rtTA^{+/+}$  (n = 4) or  $rtTA^{-/-}$  (n = 3) mice. cDNA was amplified using probes for the indicated genes or *Gapdh* (internal control). Data are means and s.d. \*\* $P = 0.001$ , \* $P = 0.04$  (*Tnf*) or 0.05 (*Il6*).

doxycycline (Dox)-inducible version of IκB-SR (see Supplementary Fig. 3a, b). KP mice were crossed to clara cell secretory protein (CCSP, also known as Scgb1a1)-rtTA transgenic mice, which express the reverse tetracycline transactivator (rtTA) specifically in lung epithelial cells<sup>23</sup>. KP; CCSP-rtTA<sup>+</sup> mice and KP; CCSP-rtTA<sup>-</sup> littermate controls were infected with TRE.IκB-SR; Pgk.Cre lentiviruses to initiate tumorigenesis. In an initial experiment, infected mice were put on a Dox diet continuously from the day of infection. After 15 weeks, lung tumour development was analysed using *in vivo* micro-computed tomography (μ-CT) imaging. As we observed with constitutive IκB-SR expression, Dox-mediated IκB-SR induction from the time of tumour initiation led to a significant impairment of tumour development (Fig. 4b). To address the role of NF-κB in tumour maintenance, cohorts of infected mice were imaged by μ-CT at 15 weeks post-infection to establish pre-treatment lung tumour volumes. One week later, the mice were placed on a Dox diet to induce the expression of IκB-SR specifically in the tumour cells, and imaged weekly for a total of 5 weeks. As shown in Fig. 4c, IκB-SR expression resulted in a significantly diminished tumour growth rate over time compared to controls. Thus, established tumours with an activating mutation in *Kras* and loss of *p53* function are sensitive to NF-κB inhibition.

To begin to explore the effects of NF-κB inhibition in this system, we examined the expression levels of certain NF-κB target genes by real-time PCR after 1 week of Dox treatment, starting at 14 weeks post-infection. Of note, *IκBα* levels (which correspond to both endogenous *IκBα* and the exogenous IκB-SR) were about fivefold higher in IκB-SR-expressing tumours, demonstrating the efficiency of the Dox-inducible system (Fig. 4d and Supplementary Table 1). Contrary to our expectations, the expression of three NF-κB targets that inhibit apoptosis, *Bcl2*, *Bclx* (also known as *Bcl2l1*) and *Xiap*, did not significantly change after Dox treatment. In contrast, *Tnf* and *Il6* messenger RNA levels were increased in tumours expressing IκB-SR. The induction of these inflammatory genes is probably due to infiltrating inflammatory cells recruited to tumours in which NF-κB has been inhibited (Supplementary Fig. 4). Such an inflammatory response has been observed after NF-κB repression in liver and intestinal epithelial cells<sup>4,24</sup>. Future experiments will examine the importance of inflammation in the anti-tumour response.

Lung cancer is the leading cause of cancer death in the world. To develop more effective treatments, it is essential to characterize the pathways that control tumour cell proliferation and survival, which will differ depending on the specific genetic context of individual tumours. Our data establish a critical role for the NF-κB pathway in lung cancer cells containing mutations in the commonly mutated genes *Kras* and *p53*. These results, which corroborate earlier studies demonstrating the importance of NF-κB signalling in a chemical-induced lung tumour model<sup>25</sup>, provide a basis to consider the pharmacological inhibition of the NF-κB pathway in the treatment of lung cancer. Because canonical NF-κB signalling has anti-inflammatory and immunosuppressive functions in macrophages<sup>26,27</sup>—phenotypes that are regularly observed in tumour-associated macrophages—NF-κB inhibition might have dual functions. Indeed, NF-κB inhibition may directly affect tumour epithelial cell survival and the tumour microenvironment, by recruiting and/or converting neutrophils or macrophages towards pro-inflammatory and tumoricidal phenotypes.

The therapeutic inhibition of NF-κB has been viewed with caution upon the discovery of diverse functions of this pathway in different physiological instances<sup>28</sup>. Our study indicates a re-evaluation of the importance of targeting NF-κB in cancer. With an a priori knowledge of the genetic mutations of the tumours, treatment with NF-κB inhibitors may become an important option for targeted therapy of lung cancer.

## METHODS SUMMARY

**Mice.**  $Kras^{LSL-G12D/WT}$ ;  $p53^{Flx/Flx}$  (KP),  $p53^{LSL/LSL}$ ; Cre-*ER*<sup>T2</sup> and  $Kras^{LA2}$  mice have been described earlier<sup>15,18,19</sup>. To generate a lung adenocarcinoma mouse model with restorable *p53*,  $p53^{LSL/LSL}$ ; Cre-*ER*<sup>T2</sup> mice were crossed to  $Kras^{LA2}$  mice (D.M.F. and T.J., unpublished data). CCSP-rtTA mice<sup>23</sup>, purchased from the Jackson Laboratory, were crossed to KP mice. A 625 doxycycline diet was from Harlan Laboratories. Mouse research was approved by the Committee for Animal Care, and conducted in compliance with the Animal Welfare Act Regulations and other Federal statutes relating to animals and experiments involving animals, and adheres to the principles set forth in the Guide for the Care and Use of Laboratory Animals, National Research Council, 1996 (institutional animal welfare assurance no. A-3125-01).

**Virus production and infection.** 293T cells were transfected using Fugene (Roche) with pMD2G (VSV-G protein), Δ8.2 (lentivirus packaging vector, provided by I. Verma) and expression plasmid. Viral supernatants were collected 48 and 72 h after transfection, filtered and used directly for infection of cell lines. For mouse experiments, virus preparation and mouse infection protocols were previously described<sup>29</sup>. Two-to-three-month-old mice were intratracheally infected with  $5 \times 10^3$ ,  $10^4$  or  $5 \times 10^4$  lentiviral particles, or with  $10^7$  adenoviral particles expressing Cre (AdCre), as indicated.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 16 April; accepted 24 August 2009.  
Published online 21 October 2009.

1. Hayden, M. S. & Ghosh, S. Signaling to NF-κB. *Genes Dev.* **18**, 2195–2224 (2004).

2. Greten, F. R. *et al.* IKK $\beta$  links inflammation and tumorigenesis in a mouse model of colitis-associated cancer. *Cell* **118**, 285–296 (2004).
3. Hanson, J. L., Hawke, N. A., Kashatus, D. & Baldwin, A. S. The nuclear factor  $\kappa$ B subunits RelA/p65 and c-Rel potentiate but are not required for Ras-induced cellular transformation. *Cancer Res.* **64**, 7248–7255 (2004).
4. Luedde, T. *et al.* Deletion of NEMO/IKK $\gamma$  in liver parenchymal cells causes steatohepatitis and hepatocellular carcinoma. *Cancer Cell* **11**, 119–132 (2007).
5. Maeda, S., Kamata, H., Luo, J. L., Leffert, H. & Karin, M. IKK $\beta$  couples hepatocyte death to cytokine-driven compensatory proliferation that promotes chemical hepatocarcinogenesis. *Cell* **121**, 977–990 (2005).
6. Pikarsky, E. *et al.* NF- $\kappa$ B functions as a tumour promoter in inflammation-associated cancer. *Nature* **431**, 461–466 (2004).
7. Herbst, R. S., Heymach, J. V. & Lippman, S. M. Lung cancer. *N. Engl. J. Med.* **359**, 1367–1380 (2008).
8. Finco, T. S. *et al.* Oncogenic Ha-Ras-induced signaling activates NF- $\kappa$ B transcriptional activity, which is required for cellular transformation. *J. Biol. Chem.* **272**, 24113–24116 (1997).
9. Huang, W. C., Ju, T. K., Hung, M. C. & Chen, C. C. Phosphorylation of CBP by IKK $\alpha$  promotes cell growth by switching the binding preference of CBP from p53 to NF- $\kappa$ B. *Mol. Cell* **26**, 75–87 (2007).
10. Kawauchi, K., Araki, K., Tobiume, K. & Tanaka, N. p53 regulates glucose metabolism through an IKK-NF- $\kappa$ B pathway and inhibits cell transformation. *Nature Cell Biol.* **10**, 611–618 (2008).
11. Kawauchi, K., Araki, K., Tobiume, K. & Tanaka, N. Loss of p53 enhances catalytic activity of IKK $\beta$  through O-linked  $\beta$ -N-acetyl glucosamine modification. *Proc. Natl Acad. Sci. USA* **106**, 3431–3436 (2009).
12. Ravi, R. *et al.* p53-mediated repression of nuclear factor- $\kappa$ B RelA via the transcriptional integrator p300. *Cancer Res.* **58**, 4531–4536 (1998).
13. Wadgaonkar, R. *et al.* CREB-binding protein is a nuclear integrator of nuclear factor- $\kappa$ B and p53 signaling. *J. Biol. Chem.* **274**, 1879–1882 (1999).
14. Webster, G. A. & Perkins, N. D. Transcriptional cross talk between NF- $\kappa$ B and p53. *Mol. Cell. Biol.* **19**, 3485–3495 (1999).
15. Jackson, E. L. *et al.* The differential effects of mutant p53 alleles on advanced murine lung cancer. *Cancer Res.* **65**, 10280–10288 (2005).
16. Jackson, E. L. *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev.* **15**, 3243–3248 (2001).
17. Jonkers, J. *et al.* Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nature Genet.* **29**, 418–425 (2001).
18. Johnson, L. *et al.* Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature* **410**, 1111–1116 (2001).
19. Ventura, A. *et al.* Restoration of p53 function leads to tumour regression *in vivo*. *Nature* **445**, 661–665 (2007).
20. Rothwarf, D. M., Zandi, E., Natoli, G. & Karin, M. IKK- $\gamma$  is an essential regulatory subunit of the I $\kappa$ B kinase complex. *Nature* **395**, 297–300 (1998).
21. Yamaoka, S. *et al.* Complementation cloning of NEMO, a component of the I $\kappa$ B kinase complex essential for NF- $\kappa$ B activation. *Cell* **93**, 1231–1240 (1998).
22. Ventura, A. *et al.* Cre-lox-regulated conditional RNA interference from transgenes. *Proc. Natl Acad. Sci. USA* **101**, 10380–10385 (2004).
23. Perl, A.-K. T., Tichelaar, J. W. & Whitsett, J. A. Conditional gene expression in the respiratory epithelium of the mouse. *Transgenic Res.* **11**, 21–29 (2002).
24. Nenci, A. *et al.* Epithelial NEMO links innate immunity to chronic intestinal inflammation. *Nature* **446**, 557–561 (2007).
25. Stathopoulos, G. T. *et al.* Epithelial NF- $\kappa$ B activation promotes urethane-induced lung carcinogenesis. *Proc. Natl Acad. Sci. USA* **104**, 18514–18519 (2007).
26. Fong, C. H. *et al.* An antiinflammatory role for IKK $\beta$  through the inhibition of “classical” macrophage activation. *J. Exp. Med.* **205**, 1269–1276 (2008).
27. Hagemann, T. *et al.* “Re-educating” tumor-associated macrophages by targeting NF- $\kappa$ B. *J. Exp. Med.* **205**, 1261–1268 (2008).
28. Pikarsky, E. & Ben-Neriah, Y. NF- $\kappa$ B inhibition: a double-edged sword in cancer? *Eur. J. Cancer* **42**, 779–784 (2006).
29. DuPage, M., Dooley, A. L. & Jacks, T. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nature Protocols* **4**, 1064–1072 (2009).
30. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative C<sub>T</sub> method. *Nature Protocols* **3**, 1101–1108 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank D. McFadden, T. Oliver, M. DuPage, K. Lane, A. Cheung, M. Kumar and E. Snyder for discussions and for sharing various reagents, A. Deconinck for critical reading of the manuscript, D. Crowley for preparation of tissue sections, and the entire Jacks laboratory for discussions. This work was supported by the Howard Hughes Medical Institute (T.J.) and partially by a Cancer Center Support grant from the NCI (P30-CA14051). T.J. is the David H. Koch Professor of Biology and a Daniel K. Ludwig Scholar. E.M. is a recipient of fellowships from the International Human Frontier Science Program Organization and the Swiss National Science Foundation. D.M.F. is a recipient of a Leukemia & Lymphoma Society Fellow Award.

**Author Contributions** E.M. and T.J. designed the study; E.M., A.L.D., L.S. and E.T. performed the experiments; D.M.F. generated the mouse lung cancer model with restorable p53; C.O. generated lung cell lines from tumours of these mice; E.M. and T.J. analysed the data; E.M. wrote the paper; and A.L.D. and T.J. edited the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to T.J. (tjacks@mit.edu).



## METHODS

**Measurement of lung tumour volumes.** Mice were anaesthetized using isoflurane, and were under anaesthesia during the entire scanning procedure. Lungs were imaged at the Koch Institute Microscopy and Imaging Core Facility by  $\mu$ -CT. Image acquisition was performed using eXplore Scan Control software, using a 45- $\mu$ m voxel size program, and three-dimensional reconstruction was performed using eXplore Reconstruction Utility software (both from GE Healthcare). High-resolution files of the lung scans were generated and individual tumour volumes were measured and calculated using MicroView software (GE Healthcare). For Fig. 4b, the lung surface opacity was set to 25%.

**Histology and immunohistochemistry.** Mouse lungs were inflated and then fixed in 3.6% formaldehyde in PBS for 24 h, transferred to 70% ethanol, and embedded in paraffin. Four-micrometre tissue sections were deparaffinized and stained with haematoxylin and eosin to count tumour numbers (Fig. 4a), or were used for immunohistochemistry. In this case, sections were deparaffinized, boiled in 10 mM Na-citrate, washed in  $H_2O$ , treated with 1%  $H_2O_2$  for 10 min, washed in  $H_2O$  and in PBS. Blocking was done for 1 h in PBS with 5% goat serum, followed by incubation with anti-p65 antibodies (C-20, 1/200) for 1 h, then by washing and staining with biotin-conjugated secondary antibodies (1 h). After washing, avidin-biotin horseradish peroxidase (HRP) complexes were added during 30 min (ABC kit, Vectastain), and the complexes were revealed with a DAB peroxidase substrate kit (Vector Laboratories). The tissues were counterstained with methyl green and mounted.

**Expression vectors.** I $\kappa$ B super-repressor (I $\kappa$ B-SR) was cloned from a mouse expressed sequenced tag with a 5' oligonucleotide that added an EcoRI site followed by a Kozak sequence before the start codon, and mutating codons for Ser 32 and Ser 36 into alanine-encoding ones. The 3' oligonucleotide introduced a silent mutation removing an EcoRI site in the gene sequence, deleted the I $\kappa$ B stop codon, added a Flag-tag sequence followed by a stop and a NotI site. I $\kappa$ B-SR was amplified using Pfx polymerase (Invitrogen), and was cloned into EcoRI-NotI sites located next to a *Ubc* promoter in a dual-promoter lentiviral vector that also expresses Cre from the *Pgk* promoter. To generate a doxycycline-responsive Flag-I $\kappa$ B-SR version, I $\kappa$ B-SR was PCR amplified with oligonucleotides adding EcoRI sites on both ends, and cloned into an EcoRI-digested TRE; *Pgk*:Cre lentiviral vector (pCW22), downstream of the TRE promoter. shRNAs against *NEMO*, *c-Rel* and *p65* were designed by the pSICOLIGOMAKER 1.5 program (<http://web.mit.edu/jacks-lab/protocols/pSico.html>; created by A. Ventura, Memorial Sloan-Kettering Cancer Center). Oligonucleotides were annealed and ligated into pSicoRev retroviral vectors, as described<sup>22</sup>. The fidelity of the PCR amplifications and oligonucleotide syntheses were confirmed by sequencing.

**Cell culture conditions.** The human embryonic kidney (HEK) 293T and mouse fibroblasts 3T3 cells were grown in DMEM. LKR (LKR10, LKR13) and KP cell lines were grown in Minimum Essential Medium (MEM). Lung cells derived from *Kras*<sup>LA2/WT</sup>; *p53*<sup>LSL/LSL</sup>; *Cre-ER*<sup>T2</sup> (D.M.F. and T.J., unpublished data) tumours were grown in DMEM. A549 and SW1573 cell lines were grown in a 1:1 mixture of F12 and DMEM. H23, H2122, H441, H460, Calu-6, H1944, H2009 and H727 cells were grown in RPMI. All human cell lines were obtained from the American Type Culture Collection (ATCC). Early passage MEFs (all obtained from the indicated genotypes in pure 129 backgrounds) were cultured with DMEM. All cell media were supplemented with 1% glutamine, 50 units ml<sup>-1</sup> penicillin, 50  $\mu$ g ml<sup>-1</sup> streptomycin and 10% heat-inactivated FBS.

**Statistics.** *P* values were determined by Student's *t*-tests.

**Nuclear-cytoplasmic fractionations.** Cells in 10-cm dishes were washed with cold PBS and collected by scraping after the addition of lysis buffer A (20 mM HEPES, pH 7.6, 20% glycerol, 10 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, 1 mM dithiothreitol (DTT), 0.1% NP-40 and a proteinase inhibitor cocktail (Roche)). Lysis was completed on ice for 10 min. Supernatants containing the cytoplasmic fractions were collected after centrifugation (400g, 4 °C, 5 min). The pellets were washed three times in lysis buffer A and then lysed in buffer B (20 mM HEPES, pH 7.6, 20% glycerol, 500 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, 1 mM DTT, 0.1% NP40 and a proteinase inhibitor cocktail (Roche)) for 30 min on ice. After centrifugation (15,000g, 4 °C, 15 min), the recovered supernatants containing the nuclear lysates were collected and frozen (−70 °C) until use. For Fig. 1a, the data are representative of two independent experiments each performed on two different MEF preparations. For Fig. 1b, the data are representative of results obtained from three independent cell lines.

**Cell viability assay.** Cells were split into 96-well plates (3,000 cells per well). After 24 h, cells were infected with the indicated retro- or lentiviruses, and 72 h later cells were fixed in methanol (10 min), dried (15 min), stained (10 min) in methylene blue dye (0.05% methylene blue in 1× borate buffer (10× borate buffer, pH 8.4, consists of 100 mM H<sub>3</sub>BO<sub>3</sub>, 25 mM Na<sub>2</sub>BO<sub>3</sub> and 120 mM NaCl)), washed three times under gentle tap water, and dried (2 h, 37 °C). After solubilization with 0.1 M HCl, cell viability was measured on a spectrophotometer (*A*<sub>650 nm</sub>). Non-infected cell values were set to 1 (100% viability).

**RNA purification, reverse transcription and real-time PCR amplification.** RNA was purified using Trizol (Invitrogen), according to the manufacturer's instructions. One-microgram RNA was reverse-transcribed using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Ten nanograms of cDNA was used for real-time PCR amplification, using commercially available Taqman probes for mouse *Gapdh*, *Tnf*, *Glut3*, *Ikbα*, *Bcl2*, *Bclx*, *Xiap*, *Il6* and *Cxcl1* (Applied Biosystems). Data were normalized to the *Gapdh* levels, and analysed using the comparative *C<sub>T</sub>* method<sup>30</sup>, except for Fig. 4d, where data are presented as fold changes ( $2^{-\Delta C_T}(\text{rtTA}^+)/2^{-\Delta C_T}(\text{rtTA}^-)$ ; mean *rtTA*<sup>−</sup> set to 1).

**NF- $\kappa$ B p65 DNA-binding activity assay.** Five micrograms of nuclear extracts was used to determine p65 DNA-binding activity using an ELISA-based assay, according to the manufacturer's instructions (Active Motif TransAM). In brief,  $\kappa$ B oligonucleotide-coated plates (in a 96-well format) were incubated for 1 h with the nuclear extracts. Specificity was achieved by incubation with anti-p65 primary antibodies for 1 h. HRP-conjugated secondary antibodies were used for the detection of p65 bound to the  $\kappa$ B sequences.

**Reagents.** Mouse TNF (used at 50 ng ml<sup>-1</sup>) was from Apotech. Antibodies to cleaved caspase-3 (no. 9661), IKK $\beta$  (2C8) and PARP (46D11) were purchased from Cell Signaling. Antibodies to p65 (C-20), c-Rel (C) and NEMO (FL-419) were purchased from Santa Cruz. Mouse anti-p53 was a gift from K. Helin. Rabbit anti-Flag (F7425) antibodies, 4-hydroxytamoxifen and doxycycline were purchased from Sigma.

**Infection of MEFs.** In Fig. 1a, primary MEFs (passage 2) of the indicated genotypes were infected with adenoviruses (100 multiplicity of infection (m.o.i.)) expressing Cre recombinase (AdCre) or FlpO recombinase (AdFlpO) (University of Iowa, Gene Transfer Vector Core, <http://www.uiowa.edu/~gene/>). Six days later, cells were collected and subjected to the cytoplasmic-nuclear fractionation protocol.

## LETTERS

# Systematic RNA interference reveals that oncogenic *KRAS*-driven cancers require *TBK1*

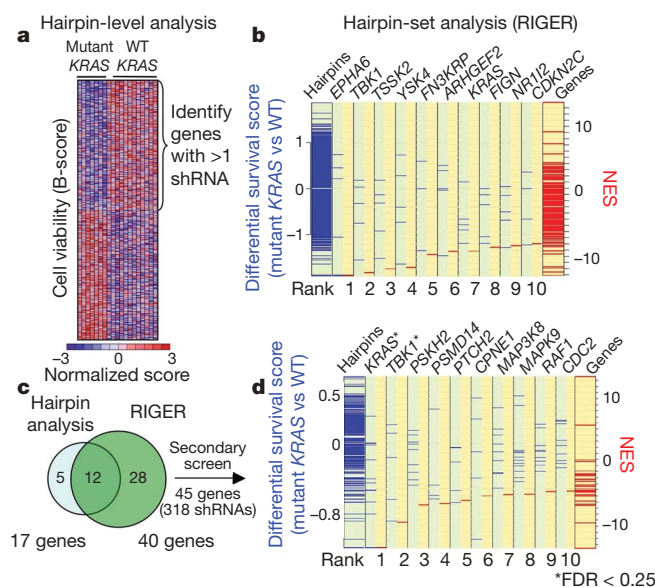
David A. Barbie<sup>1,3,4</sup>, Pablo Tamayo<sup>3</sup>, Jesse S. Boehm<sup>3</sup>, So Young Kim<sup>1,2</sup>, Susan E. Moody<sup>1,3</sup>, Ian F. Dunn<sup>1,3,5</sup>, Anna C. Schinzel<sup>1,3</sup>, Peter Sandy<sup>7,8</sup>, Etienne Meylan<sup>7,8</sup>, Claudia Scholl<sup>6</sup>, Stefan Fröhling<sup>6</sup>, Edmond M. Chan<sup>3</sup>, Martin L. Sos<sup>9</sup>, Kathrin Michel<sup>9</sup>, Craig Mermel<sup>1,3</sup>, Serena J. Silver<sup>3</sup>, Barbara A. Weir<sup>3</sup>, Jan H. Reiling<sup>7,10</sup>, Qing Sheng<sup>1</sup>, Piyush B. Gupta<sup>3</sup>, Raymond C. Wadlow<sup>3,4</sup>, Hanh Le<sup>3</sup>, Sebastian Hoersch<sup>8</sup>, Ben S. Wittner<sup>3,4</sup>, Sridhar Ramaswamy<sup>3,4</sup>, David M. Livingston<sup>1</sup>, David M. Sabatini<sup>3,7,10,11</sup>, Matthew Meyerson<sup>1,2,3</sup>, Roman K. Thomas<sup>9,12,13</sup>, Eric S. Lander<sup>3,7</sup>, Jill P. Mesirov<sup>3</sup>, David E. Root<sup>3</sup>, D. Gary Gilliland<sup>1,3,6,11</sup>, Tyler Jacks<sup>3,7,8,11</sup> & William C. Hahn<sup>1,2,3,6</sup>

The proto-oncogene *KRAS* is mutated in a wide array of human cancers, most of which are aggressive and respond poorly to standard therapies. Although the identification of specific oncogenes has led to the development of clinically effective, molecularly targeted therapies in some cases, *KRAS* has remained refractory to this approach. A complementary strategy for targeting *KRAS* is to identify gene products that, when inhibited, result in cell death only in the presence of an oncogenic allele<sup>1,2</sup>. Here we have used systematic RNA interference to detect synthetic lethal partners of oncogenic *KRAS* and found that the non-canonical I $\kappa$ B kinase *TBK1* was selectively essential in cells that contain mutant *KRAS*. Suppression of *TBK1* induced apoptosis specifically in human cancer cell lines that depend on oncogenic *KRAS* expression. In these cells, *TBK1* activated NF- $\kappa$ B anti-apoptotic signals involving c-Rel and BCL-XL (also known as BCL2L1) that were essential for survival, providing mechanistic insights into this synthetic lethal interaction. These observations indicate that *TBK1* and NF- $\kappa$ B signalling are essential in *KRAS* mutant tumours, and establish a general approach for the rational identification of co-dependent pathways in cancer.

To identify essential genes in human malignant and non-transformed cell lines, we performed arrayed format RNA interference (RNAi) screens in 19 cell lines (Supplementary Table 1 and Supplementary Fig. 1) using a short hairpin RNA (shRNA) library targeting kinases, phosphatases and oncogenes<sup>3</sup>. We then used two methods to find genes that were selectively required in cells expressing oncogenic *KRAS*. First, we used a class-discrimination feature selection method (Fig. 1a) in which normalized B-scores<sup>4</sup> for each cell line were analysed using a *t*-test statistic<sup>5</sup> to identify the top 250 (5%) shRNAs that distinguished between cell lines that contained mutant or wild-type *KRAS*. We focused on genes in which suppression by at least two shRNAs selectively impaired the proliferation/viability of *KRAS* mutant cells and identified *KRAS* itself (Supplementary Tables 2 and 3 and Supplementary Fig. 2a, b).

In parallel, we used RNAi gene enrichment ranking (RIGER)<sup>6</sup>—a statistical approach that does not rely on arbitrary thresholds—to rank-order candidate *KRAS* synthetic lethal genes (Fig. 1b). RIGER considers all shRNAs for a gene as a ‘hairpin set’, similar to ‘gene sets’ in gene set enrichment analysis (GSEA)<sup>7</sup>, and provides a normalized

enrichment score (NES) for each gene with respect to a specific classification. Using the mutant versus wild-type *KRAS* class distinction as the classification feature, we ranked candidate *KRAS* synthetic lethal partners by NES and selected the top 40 genes, which included 12 of the 17 candidates identified by the individual shRNA-based analysis (Fig. 1b, c and Supplementary Tables 2–4).



**Figure 1 | Meta-analysis of RNAi screens identifying *KRAS* synthetic lethals.** **a**, Supervised analysis of viability data (B-score) identified 250 shRNAs that distinguished mutant *KRAS* from wild-type (WT) cells, including genes targeted by several shRNAs. **b**, Hairpin-set analysis (RIGER). Genes were assigned NESs (red lines) on the basis of *KRAS* mutant/wild-type differential survival scores (blue lines) for each shRNA. Negative values represent mutant *KRAS*-selectivity. **c**, Union of 17 genes identified in **a** and 40 genes identified in **b**. **d**, Secondary screening data normalized using the percentage of control and analysed using RIGER. The FDR values for *KRAS* and *TBK1* were 0.04 and 0.18, respectively.

<sup>1</sup>Department of Medical Oncology, <sup>2</sup>Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115 USA. <sup>3</sup>Broad Institute of Harvard and M.I.T., 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. <sup>4</sup>Massachusetts General Hospital Cancer Center, 55 Fruit Street, Boston, Massachusetts 02114, USA. <sup>5</sup>Department of Neurosurgery, <sup>6</sup>Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>7</sup>Department of Biology, M.I.T., 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. <sup>8</sup>Koch Institute for Integrative Cancer Research, 40 Ames Street, Cambridge, Massachusetts 02142, USA. <sup>9</sup>Max Planck Institute for Neurological Research with Klaus-Joachim Zülch Laboratories of the Max Planck Society and the Medical Faculty of the University of Köln, Gleueler Str. 50, 50931 Köln, Germany. <sup>10</sup>Whitehead Institute of Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA. <sup>11</sup>Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA. <sup>12</sup>Department I of Internal Medicine and Center of Integrated Oncology, University of Köln, Gleueler Str. 50, 50931 Köln, Germany. <sup>13</sup>Chemical Genomics Center of the Max-Planck-Society, Otto-Hahn-Str. 15, 44227 Dortmund, Germany.

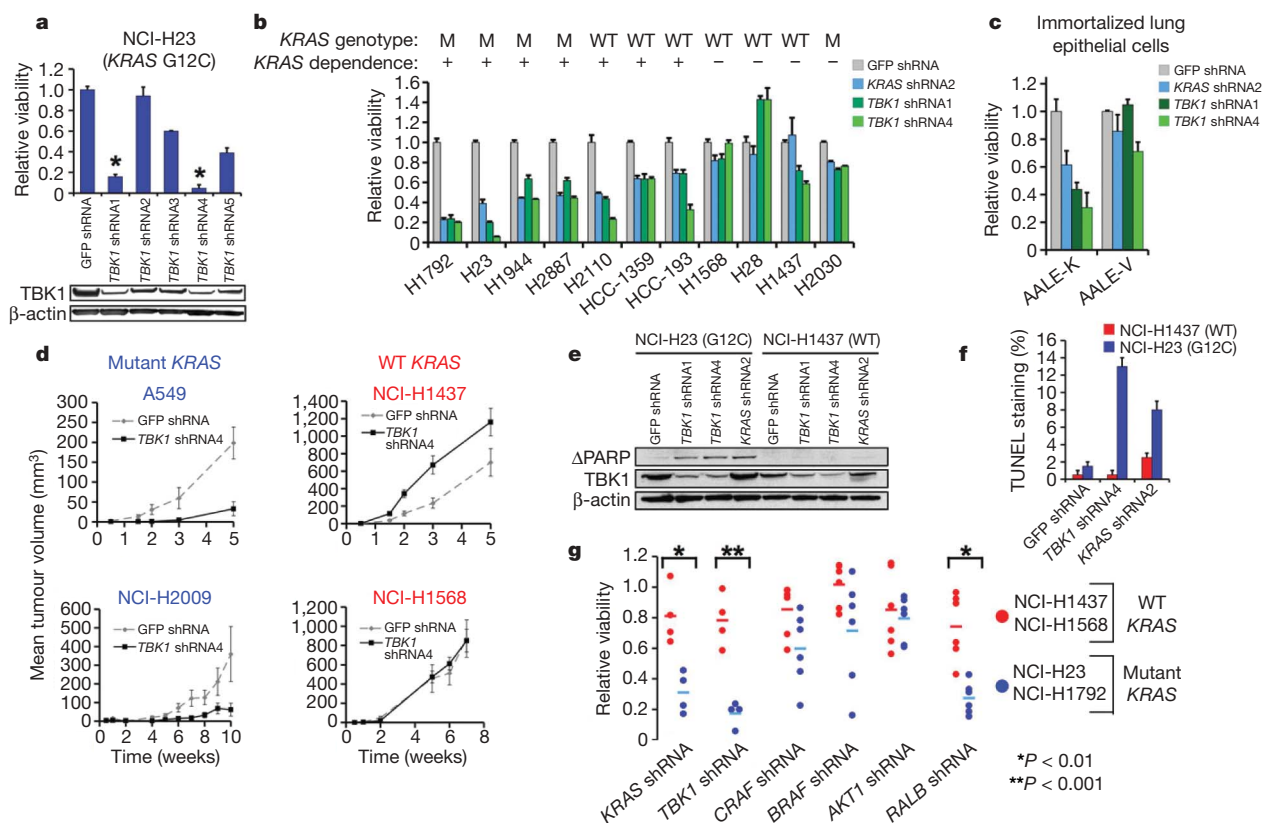
To validate the 45 candidates identified by these two approaches, we performed a secondary screen on an independent panel of mutant or wild-type *KRAS* lung adenocarcinoma cell lines (Supplementary Figs 3a, b and 4a, b). Proliferation/viability data for each shRNA was normalized to the median value of 20 control shRNAs. Using the *t*-test statistic to rank shRNAs that selectively impaired proliferation/viability in mutant *KRAS* cells, we identified a significantly enriched subset of candidate shRNAs ( $P \leq 0.0002$ ) (Supplementary Fig. 5a). Three *KRAS*-specific shRNAs were among the top four shRNAs that distinguished between *KRAS* mutant and wild-type cell lines (Supplementary Figs 4a, c and 5b). Using RIGER to rank candidate genes with respect to *KRAS*-selective lethality, we identified *KRAS* and *TBK1* as the most significant genes (false discovery rate (FDR) of 0.04 and 0.18, respectively) (Fig. 1d). Although the secondary screen identified other potential *KRAS* synthetic lethal genes, we focused on *TBK1* because it represented the top candidate after *KRAS*.

Indeed, we found that the two top-scoring shRNAs induced *TBK1* suppression and substantial cell death in NCI-H23 cells (mutant *KRAS*) (Fig. 2a). To confirm these findings, we introduced *KRAS*- or *TBK1*-specific shRNAs into a third set of lung cancer cell lines (Fig. 2b), and observed a strong correlation between *KRAS* and *TBK1* dependence, even in cell lines in which *KRAS* mutation status and dependence were decoupled. We also used an isogenic experimental model to isolate the genetic interaction between oncogenic *KRAS* and *TBK1*. Specifically, the expression of oncogenic *KRAS* in immortalized human lung epithelial cells (AALE-K cells)<sup>8</sup> rendered them dependent on both *KRAS* and *TBK1* for survival, as compared to cells expressing a control vector (AALE-V cells) (Fig. 2c). When we

suppressed *TBK1* in A549 or NCI-H2009 (mutant *KRAS*) cells, tumour formation was inhibited, whereas the suppression of *TBK1* had no effect on the tumorigenicity of NCI-H1437 or NCI-H1568 (wild-type *KRAS*) cells (Fig. 2d). These observations confirm that cancer cell lines that depend on oncogenic *KRAS* require *TBK1* expression.

To determine whether the suppression of *TBK1* in *KRAS*-dependent cells induced apoptosis, we found that, similar to shRNAs targeting *KRAS* itself (Supplementary Fig. 4b), *TBK1*-specific shRNAs provoked an increase in PARP cleavage (Fig. 2e) and TdT-mediated dUTP nick end labelling (TUNEL)-positive nuclei ( $P < 0.01$ ) (Fig. 2f and Supplementary Fig. 6a) in NCI-H23 cells (mutant *KRAS*) but not in NCI-H1437 cells (wild-type *KRAS*). Suppression of mouse *Tbk1* in cells derived from a *KRAS*-driven murine model of lung cancer (LKR-13 cells)<sup>9</sup> also induced apoptosis (Supplementary Fig. 6b).

*KRAS* activates several signalling pathways including those regulated by RAF kinases, phosphatidylinositol-3-OH kinases (PI(3)Ks) and RalGEFs (also known as RALGDSs). We found that the suppression of *RAF1*, *BRAF* or *AKT1* failed to kill *KRAS*-dependent lung cancer cell lines selectively (Fig. 2g and Supplementary Fig. 6c). *TBK1* suppression also failed to alter phosphorylated-p42/p44 MAP (mitogen-activated protein) kinase or phosphorylated-AKT levels (Supplementary Fig. 6d). In contrast, the suppression of *RALB* resulted in significant selective lethality in *KRAS*-dependent cell lines ( $P < 0.01$ , Fig. 2g). Consistent with previous work linking *RALB* with *TBK1* activation in the setting of tumour cell survival<sup>10</sup>, this observation suggested that *RALB*-*TBK1* signalling was required in cells that depend on oncogenic *KRAS*.



**Figure 2 | *TBK1* synthetic lethality with oncogenic *KRAS*.** **a**, Top-scoring *TBK1*-specific shRNAs (asterisks) induced lethality and *TBK1* suppression (immunoblot) in NCI-H23 cells (mutant *KRAS*). **b**, Suppression of *KRAS* or *TBK1* in non-small-cell lung cancer (NSCLC) cell lines. HCC-1359 and HCC-193 cells expressed RAS and NF- $\kappa$ B signatures. M, mutant. **c**, *KRAS* and *TBK1* dependence of lung epithelial cells expressing oncogenic *KRAS* (AALE-K) or vector (AALE-V). **d**, Tumour formation after *TBK1* suppression. The mean and s.e.m. of at least 11 replicates are shown.

**e**, Immunoblot of cleaved PARP ( $\Delta$ PARP) after *TBK1* or *KRAS* suppression. **f**, The percentage of TUNEL-positive nuclei after *TBK1* or *KRAS* suppression. Mean and s.d. shown. **g**, Differential cell viability after *KRAS*, *TBK1*, *CRAF*, *BRAF*, *AKT1* or *RALB* suppression using multiple shRNAs in *KRAS* mutant versus wild-type cell lines. \* $P < 0.01$ , \*\* $P < 0.001$  (*t*-test for comparisons). The s.e.m. of triplicate samples normalized to green fluorescent protein (GFP) shRNA control vector are shown.



TBK1 is a non-canonical I $\kappa$ B kinase that regulates innate immunity through the interferon and NF- $\kappa$ B pathways<sup>11</sup>, and is also a component of the exocyst complex<sup>10</sup>. To examine how TBK1 contributes to survival in *KRAS*-dependent cell lines, we performed transcriptional profiling on AALE cells expressing a control vector (AALE-V), oncogenic *KRAS* (AALE-K) or wild-type *KRAS* (AALE-K wild type). Using GSEA to identify gene sets from the Molecular Signatures Database (MSigDB-C2 v2)<sup>7</sup> that were enriched in AALE-K cells, we identified a previously described oncogenic RAS signature<sup>12</sup> as well as several NF- $\kappa$ B pathway activation signatures<sup>13,14</sup> among the most significantly enriched gene sets ( $P \leq 4.5 \times 10^{-7}$ , hypergeometric test) (Fig. 3a and Supplementary Fig. 7a). In contrast, we failed to detect enrichment of oncogenic RAS or NF- $\kappa$ B signatures in AALE-K wild-type cells (Fig. 3a), indicating that the expression of oncogenic but not wild-type *KRAS* correlated with NF- $\kappa$ B signalling.

To extend these observations to patient-derived tumours, we analysed expression profiles from 128 lung adenocarcinomas<sup>15,16</sup> for expression of the oncogenic RAS<sup>12</sup>, NF- $\kappa$ B<sup>13,14</sup> and IRF3 (ref. 17) signatures as well as a *KRAS*-specific signature (AALE-K) composed of the genes most significantly induced in AALE-K relative to AALE-V cells. We found that most mutant *KRAS* tumours (14 out of 19) showed RAS signature activation and co-expression of the NF- $\kappa$ B signature ( $P \leq 1.3 \times 10^{-15}$ , Spearman correlation test with Bonferroni adjustment) or the IKK $\epsilon$ -regulated NF- $\kappa$ B gene subset ( $P \leq 0.008$ ), but not

the IRF3-regulated gene set ( $P \leq 0.18$ ) (Fig. 3b). These observations confirm that most lung cancers that contain mutant *KRAS* show evidence of RAS and NF- $\kappa$ B pathway activation, and suggest that a substantial fraction of *KRAS* mutant primary lung cancers may depend on *TBK1* and NF- $\kappa$ B signalling for survival.

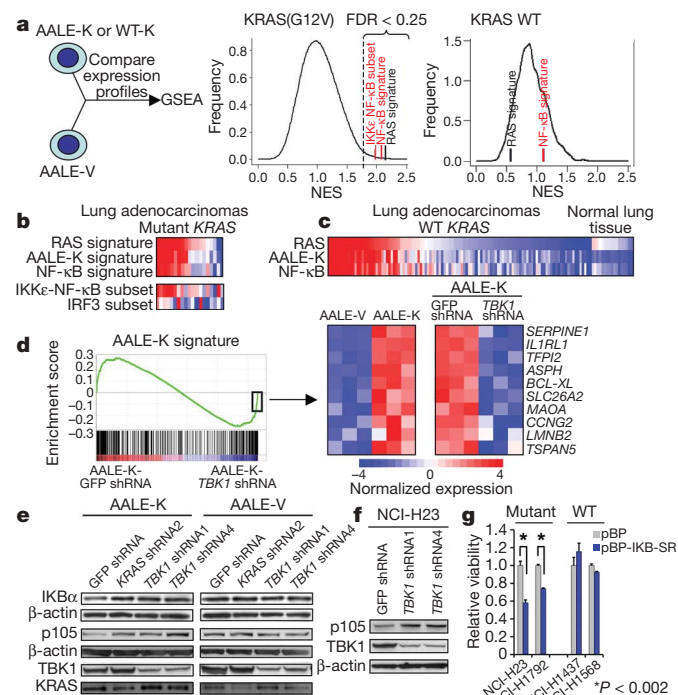
Consistent with recent work<sup>18</sup>, we also identified RAS and NF- $\kappa$ B signature co-activation in 30 out of 109 *KRAS* wild-type tumours (Fig. 3c). These RAS and NF- $\kappa$ B signatures identified some but not all of the *KRAS* wild-type cell lines that showed *KRAS* dependence (Supplementary Table 1 and Fig. 2b), suggesting that a subset of *KRAS* wild-type tumours depend on TBK1 and NF- $\kappa$ B signalling for survival. Further work will be necessary to determine whether such signatures will prove useful in predicting responsiveness to *TBK1* inhibition.

Although TBK1 activates the interferon pathway through the regulation of IRF3 and IRF7 (refs 10, 11), we failed to observe increased expression of IRF3 target genes<sup>17</sup> (Supplementary Fig. 7b) or increased IRF3 nuclear translocation (Supplementary Fig. 8a) in AALE-K cells. In addition, the suppression of *KRAS* or *TBK1* in *KRAS* mutant cancer cells downregulated specific genes within the NF- $\kappa$ B subset, including *CCND1*, *BCL2* and *IL8*, but failed to alter the expression of known interferon-responsive genes, such as *IFNB1* and *RANTES* (also known as *CCL5*) (Supplementary Fig. 7c). When we suppressed *TBK1* in AALE-K cells, we observed that NF- $\kappa$ B signature components and several NF- $\kappa$ B targets, including the anti-apoptotic gene *BCL-XL*, were among the most significantly downregulated genes (Fig. 3d). These findings confirm the importance of the NF- $\kappa$ B pathway in promoting survival in the setting of oncogenic RAS<sup>19</sup>, and suggest that, distinct from its role in innate immunity, TBK1 preferentially activates NF- $\kappa$ B signalling in tumours dependent on oncogenic *KRAS*.

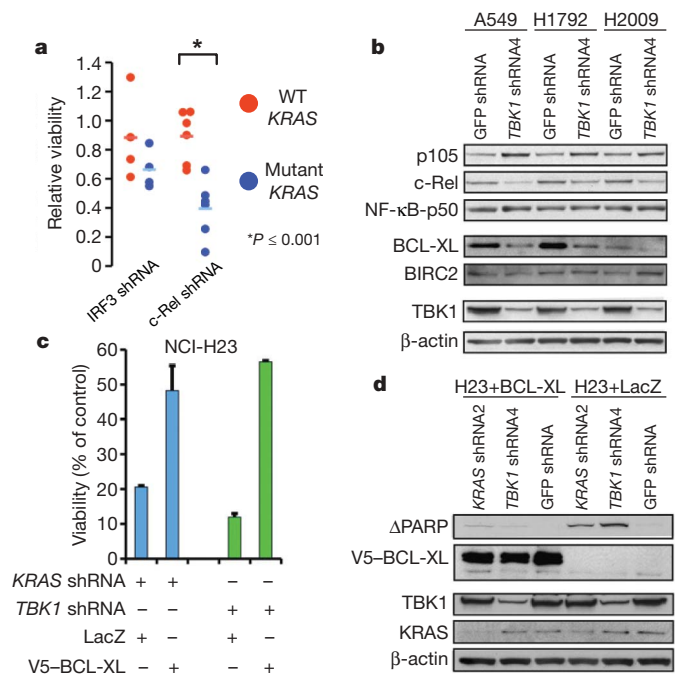
TBK1 has been reported to regulate the stability of I $\kappa$ B proteins<sup>11</sup>. When we examined cytoplasmic levels of I $\kappa$ B family members in AALE-K cells, we found reduced levels of I $\kappa$ B $\alpha$  (encoded by *NFKBIA*) and p105 (encoded by *NFKB1*) as compared to AALE-V cells (Fig. 3e and Supplementary Fig. 8a, b). Moreover, the suppression of *TBK1* in AALE-K cells or *KRAS*-mutant NCI-H23 cells (Fig. 3e, f) returned levels to that observed in wild-type *KRAS* cells. The expression of the I $\kappa$ B $\alpha$  super-repressor (IKB-SR)<sup>20</sup>, which inhibits NF- $\kappa$ B activity, in AALE-K, AALE-V or cancer cell lines expressing mutant or wild-type *KRAS* induced cell death specifically in cells containing mutant *KRAS* (Fig. 3g and Supplementary Fig. 9). These findings confirm that TBK1-driven NF- $\kappa$ B activity promotes the survival of cells that depend on mutant *KRAS*.

In the primary shRNA screen, we noted that one shRNA targeting the NF- $\kappa$ B family member *c-Rel* (also known as *REL*) scored as selectively lethal in *KRAS* mutant cells, albeit just below our pre-determined threshold. Suppression of *c-Rel* but not *IRF3* selectively induced apoptosis in *KRAS* mutant cells ( $P \leq 0.001$ ) (Fig. 4a and Supplementary Fig. 10a). Moreover, we found that the suppression of *TBK1* in *KRAS* mutant cancer cells reduced the total and nuclear *c-Rel* levels (Fig. 4b and Supplementary Fig. 10b). Although TBK1 can phosphorylate *c-Rel* when overexpressed<sup>21</sup>, we failed to detect an interaction between TBK1 and *c-Rel*, but confirmed that endogenous *c-Rel* and p105 interact (Supplementary Fig. 10c)<sup>22</sup>.

Because *BCL-XL*, a known *c-Rel* target<sup>23</sup>, was identified as a TBK1-regulated gene in AALE-K cells (Fig. 3d), we examined the expression of several anti-apoptotic genes after *TBK1* suppression in *KRAS*-mutant cancer cells, and observed specific downregulation of *BCL-XL* in several cell lines (Fig. 4b and Supplementary Fig. 10d). Moreover, the overexpression of *BCL-XL* rescued apoptosis induced by *KRAS* or *TBK1* suppression in NCI-H23 cells (Fig. 4c, d) but did not significantly affect cell death induced by the suppression of *BIRC5* (also known as *survivin*) (Supplementary Fig. 11), confirming p105, *c-Rel* and *BCL-XL* as mediators of NF- $\kappa$ B survival signalling downstream of TBK1 and *KRAS*.



**Figure 3 | Oncogenic *KRAS*-induced NF- $\kappa$ B signalling involves TBK1.** **a**, GSEA of AALE-V (vector), AALE-K (*KRAS*(G12V)) or AALE-WT-K (*KRAS* wild-type) cells (positive NES). A RAS oncogenic signature (black) and NF- $\kappa$ B signatures (red) showed significant enrichment ( $FDR < 0.25$ ) in AALE-K cells. **b**, RAS signatures in mutant *KRAS* lung adenocarcinomas correlate with NF- $\kappa$ B but not IRF3 signatures (red denotes activation, blue denotes inactivation). **c**, RAS and NF- $\kappa$ B signature expression in wild-type *KRAS* lung adenocarcinomas and normal lung tissue. **d**, AALE-K signature enrichment plot after *TBK1* shRNA or GFP shRNA expression in triplicate samples. Heatmap shows the top *KRAS*-induced genes with negative enrichment in AALE-K-*TBK1* shRNA samples. **e**, **f**, Immunoblot of I $\kappa$ B $\alpha$ , p105, TBK1 and *KRAS* in AALE-K and AALE-V cells (**e**) or NCI-H23 cells (**f**) after *KRAS* or *TBK1* suppression. **g**, Cell viability after expression of control vector (pBP) or I $\kappa$ B $\alpha$  super-repressor (pBP-I $\kappa$ B-SR) in mutant or wild-type *KRAS* cells. The mean and s.e.m. of triplicate samples are shown, and a *t*-test was used for comparisons. \* $P < 0.002$ .



**Figure 4 | TBK1 regulates c-Rel and BCL-XL in KRAS mutant cells.**

**a**, Differential cell viability after *IRF3* or *c-Rel* suppression using multiple shRNAs in *KRAS* mutant versus wild-type cell lines. **b**, Immunoblot of p105, NF-κB-p50 (cleavage product of p105), c-Rel, BCL-XL and BIRC2 in *KRAS* mutant cell lines after *TBK1* suppression. **c**, Cell viability after *KRAS* or *TBK1* suppression in NCI-H23 cells expressing a control protein (LacZ) or V5-tagged BCL-XL. The s.e.m. of triplicate samples normalized to GFP shRNA control vector are shown. **d**, Immunoblot showing overexpression of V5-tagged BCL-XL and inhibition of PARP cleavage.

We have identified *TBK1* as a synthetic lethal partner of oncogenic *KRAS*. These findings link RALB-mediated activation of *TBK1* (ref. 10) to the generation of specific NF-κB-regulated survival signals downstream of oncogenic *KRAS*. Furthermore, although studies testing the effects of inhibiting *TBK1* or NF-κB signalling in established tumours are necessary, *TBK1* and more generally NF-κB signalling may represent an alternative method of targeting oncogenic *KRAS*-driven cancers. Recently, *STK33* and *PLK1* were identified as *KRAS* synthetic lethal partners by the application of RNAi screening in paired *KRAS* mutant and wild-type cell lines<sup>24,25</sup>. Both genes were also identified in our computational analyses (Supplementary Tables 3 and 4), but like *c-Rel*, they fell below our initial threshold for secondary screening. We anticipate that the development of fully validated shRNA libraries, coupled with the interrogation of larger numbers of cell lines, will permit saturating genetic screens to identify synthetic lethal partners of *KRAS* as well as other oncogenes and tumour suppressor genes. More generally, this and other studies<sup>6,24–28</sup> indicate that the application of these functional and analytical approaches will facilitate the comprehensive identification of functional co-dependencies in cancer.

## METHODS SUMMARY

Large-scale, arrayed format RNAi screens to identify genes essential for proliferation/viability were performed as described<sup>3,14</sup>. The effect of introducing each of the 5,002 shRNAs (targeting 957 genes) was determined in 19 cell lines, and normalized using the B-score metric<sup>4</sup>. Feature selection of shRNA B-score data was performed using the Comparative Marker Application Suite in GenePattern<sup>5</sup> and was independently analysed using RIGER analysis<sup>6</sup> to compute NESs for each gene. Secondary screen viability data was normalized using a percentage of the control statistic, given the biased nature of the candidate shRNA plate. Expression profiling was used to generate a signature that correlates with *KRAS* activation and implicated NF-κB signalling in cell lines and tumours dependent on *KRAS*. Regulation of NF-κB by *TBK1* was shown using biochemical and cell biological approaches. Details of the analytical methods are provided in the Methods.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 23 January; accepted 27 August 2009.

Published online 21 October 2009.

- Hartwell, L. H., Szankasi, P., Roberts, C. J., Murray, A. W. & Friend, S. H. Integrating genetic approaches into the discovery of anticancer drugs. *Science* **278**, 1064–1068 (1997).
- Kaelin, W. G. Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nature Rev. Cancer* **5**, 689–698 (2005).
- Moffat, J. et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124**, 1283–1298 (2006).
- Malo, N., Hanley, J. A., Cerquozzi, S., Pelletier, J. & Nadon, R. Statistical practice in high-throughput screening data analysis. *Nature Biotechnol.* **24**, 167–175 (2006).
- Gould, J., Getz, G., Monti, S., Reich, M. & Mesirov, J. P. Comparative gene marker selection suite. *Bioinformatics* **22**, 1924–1925 (2006).
- Luo, B. et al. Highly parallel identification of essential genes in cancer cells. *Proc. Natl Acad. Sci. USA* **105**, 20380–20385 (2008).
- Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
- Lundberg, A. S. et al. Immortalization and transformation of primary human airway epithelial cells by gene transfer. *Oncogene* **21**, 4577–4586 (2002).
- Wislez, M. et al. Inhibition of mammalian target of rapamycin reverses alveolar epithelial neoplasia induced by oncogenic K-ras. *Cancer Res.* **65**, 3226–3235 (2005).
- Chien, Y. et al. RalB GTPase-mediated activation of the IκB family kinase TBK1 couples innate immune signaling to tumor cell survival. *Cell* **127**, 157–170 (2006).
- Häcker, H. & Karin, M. Regulation and function of IKK and IKK-related kinases. *Sci. STKE* **2006**, re13 (2006).
- Bild, A. H. et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* **439**, 353–357 (2006).
- Hinata, K., Gervin, A. M., Jennifer Zhang, Y. & Khavari, P. A. Divergent gene regulation and growth effects by NF-κB in epithelial and mesenchymal cells of human skin. *Oncogene* **22**, 1955–1964 (2003).
- Boehm, J. S. et al. Integrative genomic approaches identify *IKBKE* as a breast cancer oncogene. *Cell* **129**, 1065–1079 (2007).
- Ding, L. et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455**, 1069–1075 (2008).
- Takeuchi, T. et al. Expression profile-defined classification of lung adenocarcinoma shows close relationship with underlying major genetic changes and clinicopathologic behaviors. *J. Clin. Oncol.* **24**, 1679–1688 (2006).
- Andersen, J., VanScoy, S., Cheng, T. F., Gomez, D. & Reich, N. C. IRF-3-dependent and augmented target genes during viral infection. *Genes Immun.* **9**, 168–175 (2008).
- Singh, A. et al. A gene expression signature associated with “K-Ras addiction” reveals regulators of EMT and tumor cell survival. *Cancer Cell* **15**, 489–500 (2009).
- Mayo, M. W. et al. Requirement of NF-κB activation to suppress p53-independent apoptosis induced by oncogenic Ras. *Science* **278**, 1812–1815 (1997).
- Brown, K., Gerstberger, S., Carlson, L., Franzoso, G. & Siebenlist, U. Control of IκB-α proteolysis by site-specific, signal-induced phosphorylation. *Science* **267**, 1485–1488 (1995).
- Harris, J. et al. Nuclear accumulation of cRel following C-terminal phosphorylation by TBK1/IKKε. *J. Immunol.* **177**, 2527–2535 (2006).
- Mercurio, F., DiDonato, J. A., Rosette, C. & Karin, M. p105 and p98 precursor proteins play an active role in NF-κB-mediated signal transduction. *Genes Dev.* **7**, 705–718 (1993).
- Owyang, A. M. et al. c-Rel is required for the protection of B cells from antigen receptor-mediated, but not Fas-mediated, apoptosis. *J. Immunol.* **167**, 4948–4956 (2001).
- Scholl, C. et al. Synthetic lethal interaction between oncogenic *KRAS* dependency and *STK33* suppression in human cancer cells. *Cell* **137**, 821–834 (2009).
- Luo, J. et al. A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell* **137**, 835–848 (2009).
- Ngo, V. N. et al. A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* **441**, 106–110 (2006).
- Rottmann, S., Wang, Y., Nasoff, M., Deveraux, Q. L. & Quon, K. C. A TRAIL receptor-dependent synthetic lethal relationship between MYC activation and GSK3β/FBW7 loss of function. *Proc. Natl Acad. Sci. USA* **102**, 15195–15200 (2005).
- Silva, J. M. et al. Profiling essential genes in human mammary cells by multiplex RNAi screening. *Science* **319**, 617–620 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This work was supported in part by grants from the US National Cancer Institute (R33 CA128625, R01 CA130988) (W.C.H.) and NIH T32 CA09172-33 (D.A.B., S.E.M.), the Starr Cancer Consortium (I1-A11; W.C.H., D.G.G.), the Susan Madden Fund and an ASCO YIA (D.A.B.), a Department of Defense Prostate Cancer Postdoctoral Fellowship (S.Y.K.), a Brain Science Foundation Fellowship (I.F.D.), the Deutsche Krebshilfe (grant 107954) (R.K.T.),

the Fritz-Thyssen-Stiftung (grant 10.08.2.175; R.K.T.) and the NGFNplus-program of the German Ministry of Science and Education (BMBF, grant 01GS08100; R.K.T.). We thank C. Yu, G. Wei and members of the Hahn laboratory for discussions. High-throughput RNAi screening was conducted at the RNAi Platform of the Broad Institute of MIT and Harvard.

**Author Contributions** D.A.B., J.S.B., S.Y.K., S.E.M. and W.C.H. designed the experiments. D.A.B. and P.T. performed computational analyses. S.Y.K., I.F.D., A.C.S., P.S., C.S., S.F., P.B.G., J.H.R., Q.S. and R.C.W. performed primary RNAi screens; S.J.S., S.H., B.S.W., C.M. and B.A.W. assisted with data analysis. D.A.B. performed secondary screen with help from H.L. S.E.M. performed tumour xenograft experiments. E.M. performed experiments with murine cell lines. D.A.B., J.S.B., E.M.C., M.L.S., K.M. and

R.K.T. performed expression-profiling experiments. S.R., D.M.L., D.M.S., E.S.L., D.G.G., T.J. and D.E.R. supervised RNAi screens; M.M. and J.P.M. supervised data analysis. D.A.B. and W.C.H. wrote the manuscript. W.C.H. coordinated all aspects of the project. All authors discussed results and edited the manuscript.

**Author Information** All microarray data are available from the Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo>) under accession codes GSE17671, GSE17672 and GSE17643. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details accompany the full-text HTML version of the paper at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to W.C.H. ([william\\_hahn@dfci.harvard.edu](mailto:william_hahn@dfci.harvard.edu)).



## METHODS

**RNAi screens.** Large-scale RNAi arrayed format screening was conducted using a subset of the Broad Institute RNAi Consortium (TRC) shRNA library targeting kinases, phosphatases and other cancer-related genes<sup>3,14</sup>. shRNA designs and protocols for high throughput lentiviral production are available at <http://www.broad.mit.edu/rnai/trc/lib>. Cells were seeded in quadruplicate 384-well plates on day 0, followed by infection with  $8 \mu\text{g ml}^{-1}$  polybrene on day 1. Puromycin selection for duplicate plates (concentration individualized per cell line, most cell lines  $2 \mu\text{g ml}^{-1}$ ) was started on day 2. An ATP-based luminescence assay (Cell-Titer Glo, Promega) was used to determine the cell number on day 6. Raw luminescence values from duplicate plates were averaged, and the ratio of puromycin-positive to -negative values was used to assess infection efficiency. Data was normalized using the B-score metric—a variant of the Z-score that uses the median absolute deviation to account for plate-to-plate variability—as well as a two-way median polish to minimize row/column effects<sup>4</sup>. After excluding shRNAs with low infection efficiency, B-score values from puromycin-positive and -negative replicates were averaged for each shRNA. shRNA B-score values were aligned for the 5,002 shRNAs tested in all 19 cell lines. The secondary screen was conducted using a percentage of control statistic instead of the B-score.

**Hairpin-level analysis.** The meta-analysis of RNAi screens was performed using complementary computational approaches. The first involved conversion of the shRNA B-score file into a .res file format for input into the Comparative Marker Selection application suite in GenePattern, along with a standard classification file to generate class distinctions<sup>5</sup>. This method uses class discrimination feature selection and ranks shRNA B-score data by the *t*-test statistic or the signal-to-noise ratio to account for the difference in means between the two classes as well as the standard deviation across samples. Specifically, the difference in mean viability scores induced by each shRNA in the *KRAS* mutant class versus the wild-type class was normalized to its standard deviation using a *t*-test, and shRNAs were ranked by the *t*-test score to determine for class selective effects. To convert shRNA data to candidate genes and to reduce the likelihood of off-target effects, the top 250 (5%) of distinguishing shRNAs in the mutant *KRAS* class were filtered to identify several shRNAs targeting the same gene. Only those genes for which at least one shRNA yielded a mean B score  $< -1$  across the *KRAS* mutant class were considered. Furthermore, shRNAs with low infection efficiency in several cell lines were excluded from the analysis.

**RIGER analysis.** Similar to gene sets for GSEA<sup>7</sup>, hairpin sets were defined as the groups of shRNAs representing a given target gene. Because the number of shRNAs in a hairpin set are relatively small compared with the number of genes, we used different input data normalization, enrichment statistics and output formats compared with GSEA. RIGER analysis consisted of the following steps:

- (1) Pre-processing: the input was the same data set of B-scores analysed using the individual shRNA-based method. The B-score values for each cell line were standardized with respect to the median and maximum absolute deviation of the set of control shRNAs (directed against GFP, LacZ, RFP and luciferase) in the same cell line. This centred and rescaled the values in a sample-specific manner according to the behaviour of the control hairpins. After normalization, values below  $-3$  and above  $3$  were truncated to  $-3$  and  $3$ , respectively, to reduce the effect of outliers, and were called normalized survival scores.

- (2) Feature ranking: after pre-processing, each shRNA was assigned a 'differential survival score', which represented the difference in means of the normalized survival scores in the two phenotypic classes (for example, *KRAS* mutant versus wild type). The difference in means was used to emphasize the absolute magnitude of the survival differences and not only the profile 'shape'. Specifically, this favoured shRNAs with strong differential killing of cells over ones that exhibited weak differential killing of cells, but had perfect discrimination profiles inside classes. The differential survival scores were computed for all the 5,002 shRNAs and sorted from high to low scores.

- (3) Calculation of enrichment scores: a given gene was assigned an enrichment score according to the distribution of differential survival scores of its shRNAs within the rank list of all shRNAs using a two-sample weighted 'Zhang C' statistic based on the likelihood ratio<sup>29</sup>. The Zhang C likelihood ratio statistic was used owing to its greater sensitivity and better empirical results in exploratory analyses with other data sets. We adapted this method to separate positive and negative enrichment contributions, and used a weighting factor based on the differential enrichment score, except for scores between  $-0.5$  and  $0.5$ , which were truncated to reduce the effects of shRNAs that weakly distinguished classes. The enrichment score obtained in this manner was representative of both the extremeness of the shRNA differential survival scores for a given gene and their consistency.

- (4) Normalization of enrichment scores: because genes with different numbers of shRNAs were assigned enrichment scores on different scales, we normalized them before sorting the genes by using a null distribution generated by 1,000 random permutations of the locations of the shRNAs in the entire list. The

normalization for negative enrichment scores was a rescaling by the absolute value of the mean of the negative values in the null distribution. This represented an effective way to place the enrichment scores on a common scale regardless of the number of shRNAs for each gene. The null distribution also provided nominal *P*-values for each gene enrichment score.

(5) Generation of results: the analysis resulted in a list of genes sorted by their NESs, and a set of complementary estimates of statistical significance, such as nominal, family-wise and Bonferroni *P*-values plus a FDR. A collection of dual-vertical plots was used to demonstrate the shRNA differential survival scores for each gene (lines in blue) and NESs (lines in red). Vertical plots were arranged starting with the top gene (strongest negative NES) on the left.

**Secondary screen.** Analysis of secondary screen data followed the same methodology except for the normalization of the cell proliferation/viability data. Twenty control shRNAs directed against RFP, LacZ and luciferase were screened in parallel with candidates. We normalized data for each shRNA in each cell line using the percentage of control statistic, dividing the raw data for each shRNA by the median of control shRNA values, and taking the  $\log_{10}$  of this number to scale values around zero. Because of the biased nature of the candidate shRNA plate, and because the number of control shRNAs was smaller in the secondary screen than the first, the calculation of the maximum absolute deviation was noisy and unreliable. No truncation was applied to the resulting values.

A larger set of 84 control shRNAs directed against GFP, RFP, LacZ and luciferase was also tested independently in all eight cell lines chosen for the secondary screen. We normalized data for each shRNA in each cell line by dividing the raw data for each control shRNA by the plate median and taking the  $\log_{10}$  of this number to scale values around zero. We used the *t*-test statistic to examine the *KRAS* mutant versus wild-type class distinction for this control plate relative to the candidate plate, restricting the analysis to shRNAs with strong effects on proliferation/viability (mutant *KRAS* class mean log percentage of control  $< -0.2$ , corresponding to  $\sim 37\%$  viability impairment) (Supplementary Fig. 5a). We used the *t*-test statistic threshold that was achieved by the set of control shRNAs as the boundary to identify the top 25 shRNAs *KRAS* synthetic lethal shRNAs that scored on the candidate plate (Supplementary Fig. 5b).

**Gene expression profiling or GSEA.** RNA was prepared from AALE cells expressing *KRAS*(G12V) or a control vector 6 days after infection and analysed using human U133A HTA Arrays (Affymetrix). GSEA was performed using gene sets from the Molecular Signatures Database (MSigDB-C2 v2)<sup>7</sup>. In brief, the method consists of the following steps: genes are first ranked in a list, *L*, by the correlation between their expression and the class distinction (for example, *KRAS* mutant versus wild type), using a suitable correlation metric. Given a defined set of genes *S* (for example, genes members of a signalling pathway, located in the same genomic region, sharing the same Gene Ontology category, and so on), the goal of GSEA is to determine whether the members of *S* are found at the top or bottom of the list, indicating that they associate with the phenotypic distinction, rather than being distributed uniformly or randomly across the list. Next, to evaluate this degree of 'enrichment', an enrichment score is calculated to quantify the degree to which a set *S* is over-represented at the top or bottom of the entire ranked list *L*. After calculation of the scores for a collection of gene sets, an empirical phenotype-based permutation test procedure is used to estimate *P*-values. The permutation of class labels preserves gene-gene correlations and provides an assessment of significance that is more reflective of the underlying biology. Finally, an adjustment is made to the estimated significance level to account for multiple hypotheses testing. GSEA normalizes the enrichment score for each gene set to account for the variation in set sizes, yielding a NES and a FDR. The FDR gives an estimate of the probability that a set with a given NES represents a false-positive finding; it is computed by comparing the tails of the observed and permutation-computed null distributions for the NES. The collection of gene sets used in the analysis of Fig. 3a consisted of release 2.5 of the C2 (curated gene sets) sub-collection of the Molecular Signatures Database (<http://www.broad.mit.edu/gsea/msigdb/>). To determine the significance of identifying many NF- $\kappa$ B upregulated gene sets enriched in AALE-K cells, we used a hypergeometric test.

Gene expression profiling was also performed in triplicate in AALE-K cells (*KRAS*(G13D)) 5 passages after stable integration of the *KRAS* allele and compared with AALE-V cells using human U133A Arrays (Affymetrix). We created a new *KRAS*-specific gene signature using the mean difference in expression between AALE-K and AALE-V triplicate samples to determine the 300 most significantly induced genes by oncogenic *KRAS*. In parallel, expression profiling was performed in AALE-K cells 48 h after expression of GFP shRNA or *TBK1* shRNA, to identify which of these genes were most significantly downregulated after *TBK1* suppression.

**Signature projection method.** The RAS oncogenic signature<sup>12</sup>, AALE-K signature, NF- $\kappa$ B signature<sup>13</sup>, IKK $\epsilon$ -regulated NF- $\kappa$ B gene set<sup>14</sup> and IRF3 target gene set<sup>17</sup> were projected across 38 lung adenocarcinomas derived from the

Tumour Sequencing Project<sup>15</sup>, an additional 90 lung adenocarcinomas<sup>16</sup>, a collection of 53 lung cancer cell lines<sup>30</sup>, and 17 normal lung tissue specimens<sup>31</sup>. This was accomplished by a 'single sample' extension of GSEA<sup>7</sup> that allows one to define an enrichment score that represents the degree of absolute enrichment of a gene set in each sample within a given data set. The gene expression values for a given sample were rank-normalized, and an enrichment score was produced using the Empirical Cumulative Distribution Functions (ECDF) of the genes in the signature and the remaining genes. This procedure is similar to GSEA but the list is ranked by absolute expression (in one sample). The enrichment score is obtained by an integration of the difference between the ECDF. For a given signature  $G$  of size  $N_G$  and single sample  $S$ , of the data set of  $N$  genes, the genes are replaced by their ranks according to their absolute expression from high to low:  $L = \{r_1, r_2, \dots, r_N\}$ . An enrichment score  $ES(G, S)$  is obtained by a sum (integration) of the difference between a weighted ECDF of the genes in the signature  $P_G^w$  and the ECDF of the remaining genes  $P_{N_G}$ :

$$ES(G, S) = \sum_{i=1}^N [P_G^w(G, S, i) - P_{N_G}(G, S, i)]$$

$$\text{where } P_G^w(G, S, i) = \sum_{r_j \in G, r_j \leq i} \frac{|r_j|^\alpha}{\sum_{r_j \in G} |r_j|^\alpha}$$

$$\text{and } P_{N_G}(G, S, i) = \sum_{r_j \notin G, r_j \leq i} \frac{1}{(N - N_G)}$$

This calculation is repeated for each signature and each sample in the data set. Note that the exponent of this quantity ( $\alpha$ ) is set to 1/4, and adds a modest weight to the rank. In the regular GSEA a similar enrichment score is used, but the weight is typically set to 1. Also, instead of the sum over  $i$ , the enrichment score is computed according to the largest difference. This quantity is slightly more robust and more sensitive to differences in the tails of the distributions than the Kolmogorov–Smirnov statistic. It is particularly well suited to represent the activation score of gene sets on the basis of a relatively small subset of the genes attaining high expression values. Signature values were normalized using the entire set of 128 lung adenocarcinomas and 17 normal lung specimens.  $P$  values were calculated for the lung adenocarcinoma samples testing the hypotheses that the Spearman correlation between the RAS oncogenic signature, NF- $\kappa$ B gene set, and IRF3 target gene set were greater than zero.

**Cell culture.** 293T and A549 cells were grown in DMEM supplemented with 10% FBS, 2 mM L-glutamine, penicillin (1,000 U ml<sup>-1</sup>) and streptomycin (1,000  $\mu$ g ml<sup>-1</sup>). NCI-H23, NCI-H28, HCC-193, NCI-H522, HCC-1359, NCI-H1437, NCI-H1568, NCI-H1792, NCI-H1944, NCI-H1975, NCI-2009, NCI-H2030, NCI-H2110 and NCI-H2887 cells were cultured in RPMI supplemented with 10% FBS, 2 mM L-glutamine, penicillin and streptomycin. AALE cells<sup>8</sup> were cultured in serum-free SABM media with SingleQuot supplements/growth factors (Lonza). AALE cells expressing KRAS(G12V) or KRAS(G13D) alleles were generated after transduction using a pBabe retroviral vector or pLenti6.2/V5-Dest lentiviral vector (Invitrogen) respectively, as described<sup>14</sup>. Cell lines expressing the IKB $\alpha$  super-repressor were generated using a pBabe retroviral vector expressing the IKB $\alpha$  super-repressor<sup>14</sup>, and NCI-H23 cells expressing BCL-XL or LacZ were generated using pLenti6.2/V5-Dest encoding either BCL-XL or LacZ. Cells were seeded in 96-well plates for cell viability assays, and in 6-well plates to prepare lysates for immunoblotting.

**Low throughput lentiviral shRNA production/infection.** Lentiviral vectors encoding shRNAs specific for control GFP sequences as well as KRAS, TBK1, CRAF, BRAF, AKT1, RALB, IRF3, c-Rel and BIRC5 are part of the TRC shRNA library. Sequences of validated shRNAs are listed in Supplementary Table 6. Lentiviruses were produced by transfection of 293T cells with vectors encoding gene-specific shRNAs (1  $\mu$ g) together with the packaging plasmids encoding  $\Delta$ 8.9 and VSV-G using Eugene 6 (Roche). Culture supernatants containing lentivirus were collected 48 and 72 h after transfection. Virus was pooled and stored at -80 °C. Cells were infected using a 1:11 dilution of virus in polybrene-containing media. After centrifugation at 1,000g for 15 min, all NSCLC lines were selected in puromycin (2  $\mu$ g ml<sup>-1</sup>) starting 24 h after infection. AALE cells were treated with virus/polybrene for 4 h and selected with puromycin

(1  $\mu$ g ml<sup>-1</sup>). Viability assays were conducted 6 days after infection using Cell-Titer Glo (Promega) in triplicate. Lysates were collected 72 h after shRNA expression to assess gene suppression. To determine differential viability effects in NCI-H23/NCI-H1792 (mutant KRAS) versus NCI-H1437/NCI-H1568 (wild-type KRAS) cells, mean viability for each shRNA was normalized to GFP shRNA control. Results from validated shRNAs were grouped together for each gene in KRAS mutant versus wild-type cell lines, and an unpaired  $t$ -test was used to determine statistical significance.

**Antibodies.** Immunoblotting was performed as described<sup>14</sup>. Antibodies were obtained from Cell Signaling Technology (anti-AKT1 no.2967, anti-phospho-AKT Ser 473 no.9271, anti-BCL-XL no.2762, anti-BIRC2 no.4952, anti-CRAF no.9422, anti-c-Rel no.4727, anti-GAPDH no.2118, anti-IBK $\alpha$  no.4814, anti-Lamin A/C no.2032, anti-phospho-MAPK p42/44 no.9102, anti-p105/p50 no.3035, anti-PARP cleaved Asp 214 no.9546, anti-RALB no.3523 and anti-survivin no.2808), Santa Cruz Technology (anti-KRAS sc-30, anti-B-RAF F-7, anti-IRF3 FL-425 and anti-IBK $\alpha$  c-15), and Upstate Biotechnologies/Millipore (anti-TBK1 clone AOW9).

**Tumorigenicity assay.** Tumour xenograft experiments were performed as described<sup>14</sup>. Control GFP or TBK1-specific shRNAs were expressed in the indicated cells for 72 h, and  $2 \times 10^6$  viable cells were then injected subcutaneously per site into immunodeficient mice. The mean and s.e.m. tumour volume were plotted over time. The total number of tumours/implantations were: A549-GFP shRNA (10/13); A549-TBK1 shRNA (0/11); NCI-H1437-GFP shRNA (12/12); NCI-H1437-TBK1 shRNA (12/12); NCI-1568-GFP shRNA (12/12); NCI-1568-TBK1 shRNA (12/12); NCI-H2009-GFP shRNA (9/12); NCI-H2009-TBK1 shRNA (3/11). Tumour determination was made at 3 weeks except for NCI-H2009, which were measured at 5 weeks.

**TUNEL assay.** NCI-H23 cells and NCI-H1437 cells were infected in 96-well plates with shRNAs specific for GFP, TBK1 and KRAS, fixed 5 days after infection using 10% paraformaldehyde, and subjected to TUNEL staining (Roche). Nuclei were co-stained with 4',6-diamidino-2-phenylindole (DAPI), and then imaged and counted using a Zeiss Axiovert 200 immunofluorescence microscope.

**Nuclear/cytoplasmic fractionation.** Cells were washed twice in ice-cold PBS and incubated on ice for 10 min after treatment with hypotonic cytoplasmic lysis buffer (20 mM HEPES, pH 7.6, 10 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, 1 mM dithiothreitol (DTT), 0.1% NP40 and 20% glycerol) plus proteinase inhibitors. Nuclei were pelleted at 400g at 4 °C for 4 min. Supernatants were collected, and the nuclear pellet was washed twice using cytoplasmic lysis buffer. Nuclear lysis buffer (same as cytoplasmic lysis buffer except 500 mM NaCl) was added to the pellet, and samples were incubated for 30 min on ice. After centrifugation at 16,000g at 4 °C for 15 min, the supernatant was collected as the nuclear fraction.

**Immunoprecipitation.** Cells were washed twice with PBS, treated with lysis buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 0.5% NP-40 and 2 mM EDTA) and rotated at 4 °C for 60 min. After centrifugation at 16,000g, lysates were quantified. Anti-c-Rel antibody (5  $\mu$ l) or buffer alone was added to lysate (750  $\mu$ g) in 500  $\mu$ l of lysis buffer, and tubes were rotated overnight at 4 °C. Fifty microlitres of 50% protein-G-sepharose beads were added, and samples were incubated for 2 h at 4 °C. After centrifugation at 800g for 3 min at 4 °C, beads were washed three times in lysis buffer. Elution was performed using 2 $\times$  sample buffer, and samples were divided in half and loaded onto parallel gels, along with beads only control and 1/10 input.

**Real-time quantitative RT-PCR.** Relative mRNA expression was determined using real-time quantitative PCR and normalized to GAPDH expression as an internal amplification control. Total RNA was isolated using Trizol reagent (Gibco), and 1  $\mu$ g of total RNA was reverse transcribed using SuperScript First-Strand Synthesis System (Invitrogen), followed by amplification using SYBR Green PCR Master Mix (Applied Biosystems).

29. Zhang, J. Powerful goodness-of-fit tests based on the likelihood ratio. *J. R. Stat. Soc. Series B Stat. Methodol.* **64**, 281–294 (2002).
30. Sos, M. L. et al. Predicting drug susceptibility of non-small cell lung cancers based on genetic lesions. *J. Clin. Invest.* **119**, 1727–1740 (2009).
31. Bhattacharjee, A. et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc. Natl Acad. Sci. USA* **98**, 13790–13795 (2001).

# Rationally tuning the reduction potential of a single cupredoxin beyond the natural range

Nicholas M. Marshall<sup>1</sup>, Dewain K. Garner<sup>1</sup>, Tiffany D. Wilson<sup>1</sup>, Yi-Gui Gao<sup>1</sup>, Howard Robinson<sup>2</sup>, Mark J. Nilges<sup>1</sup> & Yi Lu<sup>1</sup>

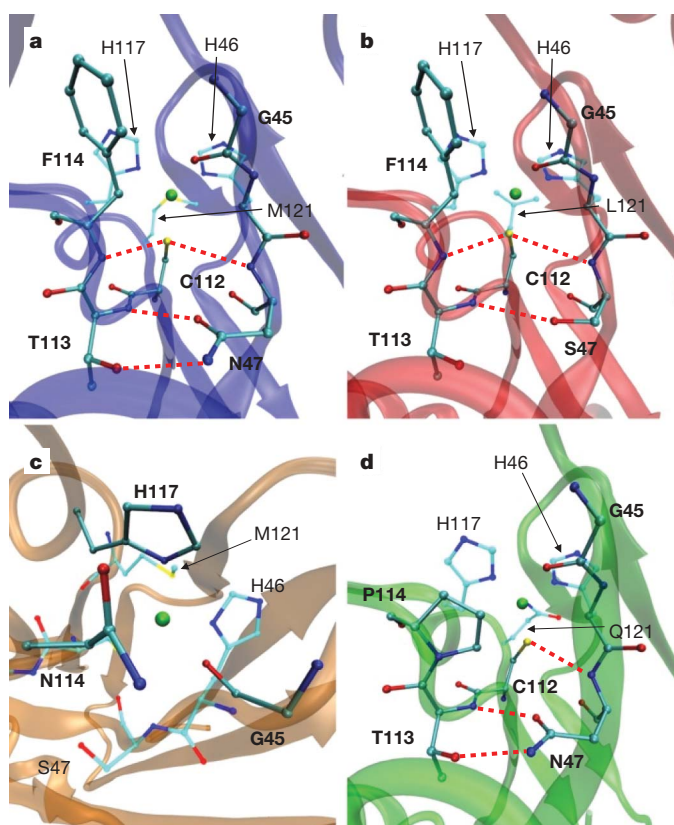
Redox processes are at the heart of numerous functions in chemistry and biology, from long-range electron transfer in photosynthesis and respiration to catalysis in industrial and fuel cell research. These functions are accomplished in nature by only a limited number of redox-active agents. A long-standing issue in these fields is how redox potentials are fine-tuned over a broad range with little change to the redox-active site or electron-transfer properties. Resolving this issue will not only advance our fundamental understanding of the roles of long-range, non-covalent interactions in redox processes, but also allow for design of redox-active proteins having tailor-made redox potentials for applications such as artificial photosynthetic centres<sup>1,2</sup> or fuel cell catalysts<sup>3</sup> for energy conversion. Here we show that two important secondary coordination sphere interactions, hydrophobicity and hydrogen-bonding, are capable of tuning the reduction potential of the cupredoxin azurin over a 700 mV range, surpassing the highest and lowest reduction potentials reported for any mononuclear cupredoxin, without perturbing the metal binding site beyond what is typical for the cupredoxin family of proteins. We also demonstrate that the effects of individual structural features are additive and that redox potential tuning of azurin is now predictable across the full range of cupredoxin potentials.

Cupredoxins are redox-active copper proteins that span several kingdoms in biology and play crucial roles in many important processes, including photosynthesis, respiration, cell signalling, and many other reactions in oxidases and reductases<sup>4–8</sup>. Despite these many functions, cupredoxins use a single redox active centre, whose reduction potential ( $E^\circ$ ) must be tunable to match that of a given redox partner without compromising the structural and electron-transfer properties of the protein. The protein environment provides not only a means to overcome the large (>1 eV) reorganization energies of Cu complexes in aqueous solution, but also the limited range of  $E^\circ$  attainable by free Cu ions under physiological conditions (<200 mV). The mononuclear type 1 (T1) cupredoxins provide a notable example, as they all contain a single Cu ion with similarly low (0.7 eV) reorganization energies, yet  $E^\circ$  values from various members of this ubiquitous family span an ~500 mV range<sup>4–8</sup>.

Despite their vastly different  $E^\circ$ , the overall and active site structures are very similar in all T1 cupredoxins<sup>9</sup>, which contain a Cu ion coordinated by a Cys and two His residues in a trigonal planar geometry, and a weak axial ligand, such as Met 121 in azurin (Az; Fig. 1a). A weak interaction between the Cu and a backbone carbonyl oxygen is also present in some cupredoxins, such as Gly45 in Az (Fig. 1a). The similar structure of various cupredoxins results in similar spectroscopy, characterized by a ligand to metal charge transfer absorbance at ~600 nm and very small electron paramagnetic resonance (EPR) hyperfine coupling constants ( $A_{||} \approx 60$  G). That the highly similar Cu-binding site structure exhibits such a wide range of  $E^\circ$  suggests

that subtle interactions outside the primary coordination sphere are used in nature to attain the required  $E^\circ$  in cupredoxins.

A number of studies have been performed to elucidate structural features responsible for controlling the  $E^\circ$  of cupredoxins<sup>4–8</sup>. Factors



**Figure 1 | X-ray structures of Az and selected variants.** **a**, Native azurin (PDB 4AZU). **b**, N47S/M121L azurin: N47S affects the rigidity of the copper binding site and, probably, the direct hydrogen bonds between the protein backbone and Cys 112. **c**, N47S/F114N azurin: introducing a hydrogen-bond donor at position 114 perturbs hydrogen-bonding near the copper binding site, possibly disrupting donor–acceptor interactions to His 117, or ionic interactions between the copper and the carbonyl oxygen of Gly 45. **d**, F114P/M121Q azurin: F114P deletes a direct hydrogen bond to Cys 112 resulting in a lower redox potential. The ultraviolet–visible spectroscopy of the F114P-containing variants shows a significant increase in the copper  $d \rightarrow d$  absorbance range around 800 nm. This increased absorbance suggests slight rearrangement of the copper binding site, but is consistent with F114P Az<sup>27</sup> and other T1 copper proteins, such as plastocyanin<sup>9</sup>. In all panels copper is shown in green, carbon in cyan, nitrogen in blue, oxygen in red and sulphur in yellow. Hydrogen-bonding interactions are shown by dashed red lines.

<sup>1</sup>Department of Chemistry, University of Illinois, Urbana-Champaign, Illinois 61801, USA. <sup>2</sup>Biology Department, Brookhaven National Laboratory, Upton, New York 11973-5000, USA.



identified in other proteins as crucial for tuning  $E^\circ$ , such as electrostatics<sup>10</sup> and solvent<sup>11,12</sup> effects, have thus far shown minimal effect in cupredoxins, suggesting that there is still a significant lack of knowledge about other long-range, non-covalent interactions responsible for tuning  $E^\circ$ . Mutational studies on rusticyanin (Rc), the mononuclear cupredoxin with the highest  $E^\circ$  of 680 mV at pH 2.0, resulted in a lowest  $E^\circ$  of ~470 mV, still higher than the highest  $E^\circ$  of other cupredoxins<sup>4–8</sup>. Similarly, numerous mutations in other cupredoxins always reach a  $E^\circ$  ceiling, with the highest  $E^\circ$  being ~400 mV, still lower than that of Rc. So far, replacement of the axial copper ligand (Met 121 in Az) has resulted in the largest changes to the  $E^\circ$  of several cupredoxins<sup>4–8</sup>. Studies using unnatural amino acids as isostructural analogues of Met established a linear correlation between the  $E^\circ$  and the hydrophobicity of the axial ligand in cupredoxin variants (Fig. 2a)<sup>13,14</sup>. Other interactions, such as  $\pi$ – $\pi$  interactions between aromatic residues and the copper ligands, have been shown to affect cupredoxin  $E^\circ$  to a lesser degree<sup>15,16</sup>. While these studies exemplify the importance of long range

(>3 Å) interactions in tuning the properties of metalloproteins, the level of change to the  $E^\circ$  is relatively small (<250 mV)<sup>15,16</sup>, and no report has shown control of the  $E^\circ$  of a single cupredoxin across the entire range of native cupredoxin potentials at similar pH values. As there are significant pH effects on the reduction potential of every cupredoxin protein, all potentials in this paper will be compared at pH 7.0, unless otherwise noted.

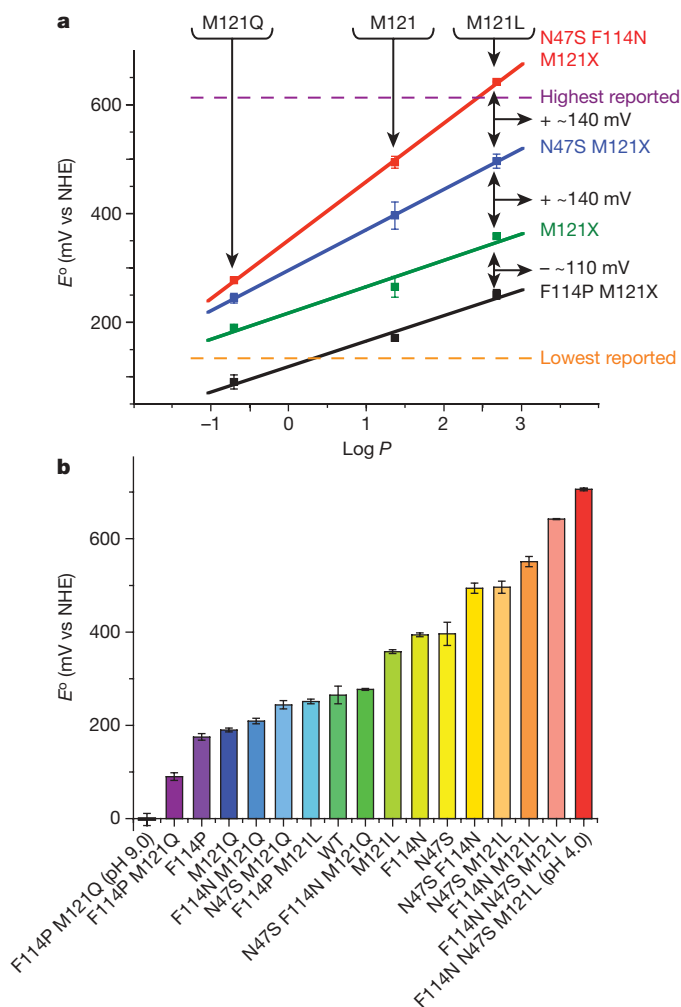
With its moderate  $E^\circ$  of  $265 \pm 19$  mV at pH 7.0, Az is an ideal system to test rational tuning of cupredoxin  $E^\circ$ , as it represents an intermediate step between the potential extremes. Inspection of several cupredoxin structures and comparison with that of Rc revealed a Ser residue (Ser86 in Rc) at a position corresponding to a highly conserved Asn in other cupredoxins. This Asn residue is ~5.5 Å away from the copper in Az, but forms hydrogen bonds between two of the ligand containing loops. Mutating Ser 86 in Rc to Asn lowered the  $E^\circ$  by 77 mV (ref. 17). Ser 86 in Rc has been proposed to raise the Cu  $E^\circ$  by strengthening the hydrogen-bonding interactions between two ligand-containing loops, adding rigidity to these loops and influencing the direct hydrogen bonds between the backbone of the protein and the thiolate of Cys ligand (refs 17–19). We therefore incorporated these interactions into Az by making the N47S mutation. The ultraviolet–visible absorption spectroscopy of N47S is very similar to that of native Az (Supplementary Fig. 3), suggesting that the core cupredoxin structure is preserved. Interestingly, the mutation resulted in an ~130 mV increase in  $E^\circ$ , from  $265 \pm 19$  mV to  $396 \pm 25$  mV at pH 7.0.

Because the axial ligand, Met 121, and the N47S mutation influence different aspects of the copper binding site, we reasoned that their individual effects on  $E^\circ$  would be additive. Consequently, we constructed both N47S/M121Q and N47S/M121L Az. The linear dependence of the redox potential on the hydrophobicity of the axial residue, previously identified in cupredoxins, was still observed, with N47S mutation exerting similar effects on their  $E^\circ$  (Fig. 2a), suggesting that the two influences are additive. Owing to this additive nature, the highest  $E^\circ$  of this series of Az mutants, seen in N47S/M121L Az, was observed to be  $496 \pm 13$  mV at pH 7.0. As shown by a crystal structure of N47S/M121L Az (Fig. 1b), the N47S mutation in Az introduces a hydroxyl group 4.13 Å from the sulphur of Cys 112 and 3.42 Å from the backbone amide nitrogen of Thr 113, reproducing a hydrogen-bonding pattern similar to that of native Rc (PDB code 1rcy)<sup>20</sup>, which shows distances of 4.04 Å and of 2.88 Å, respectively, at the equivalent positions.

Despite N47S/M121L Az having the highest redox potential of any known Az variant, its  $E^\circ$  still falls short of that of Rc. Therefore, other factors that tune the cupredoxins  $E^\circ$  to high values must still be missing. Another examination of cupredoxin crystal structures reveals the presence of a backbone carbonyl oxygen, from Gly 45, near the copper ion in Az, which is missing in other cupredoxins, such as Rc (refs 6, 7). This ionic interaction in Az results in more electron density near the copper, preferentially stabilizing the Cu(II) form of the protein and, therefore, lowering the redox potential<sup>21</sup>. We thus reasoned that manipulating the electronic properties of the carbonyl oxygen in Az may enable us to close the gap between the high and low potential cupredoxins.

As the carbonyl oxygen itself cannot be changed by site-directed mutagenesis, altering its properties represents a new challenge for protein design and engineering. To meet this challenge, we searched for residues in Az that could affect the carbonyl through either steric repulsion or hydrogen bonding. Phe 114 was chosen as a suitable candidate because of its proximity to Gly 45 and since its side chain points towards the carbonyl. To observe the effect of introducing a hydrogen-bonding group at position 114 in Az, F114N Az was constructed and a potential of  $398 \pm 4$  mV at pH 7.0 was measured for this variant.

Because both the N47S and F114N mutations individually perturb separate areas of the protein, they were combined by constructing the N47S/F114N variant of Az. The redox potential of N47S/F114N Az



**Figure 2 | Rational tuning of the reduction potential of Az.** **a**, Plot of  $E^\circ$  at pH 7.0 versus log  $P$  for the Az mutants from this study (here  $P$  is the partition coefficient of the side chain of the residue between octanol and water; that is, it is a measure of hydrophobicity). The lowest reported  $E^\circ$  (dashed orange line) at pH 7.0 for any T1 cupredoxin before this study, M86Q pseudoazurin, is indicated. The highest  $E^\circ$  (dashed purple line) at pH 6.2 of any T1 cupredoxin variant before this study, M148L Rc, is also indicated; its potential was not measured at pH 7.0 owing to protein instability. Considering the pH trend of the  $E^\circ$  of all cupredoxins, this reported value for M148L Rc would be lower at pH 7.0. **b**, Plot showing the  $E^\circ$  value for each azurin variant at pH 7.0 unless otherwise noted. Not only is the redox potential of Az tunable to the extremes of the redox potentials attainable by T1 cupredoxins, but to nearly any redox potential within the range. Error bars,  $\pm 1$  s.d.

was observed to be  $494 \pm 11$  mV at pH 7.0 (Fig. 2). Again, the position and extinction coefficient of the ligand to metal charge transfer band in this double mutant are similar to wild-type Az (Supplementary Fig. 5). Combining the axial Met mutations with N47S and F114N also yielded a linear dependence of reduction potential on axial ligand hydrophobicity, resulting in a reduction potential for N47S/F114N/M121L Az of  $640 \pm 1$  mV at pH 7.0 (Fig. 2). The redox potential of Rc at the same pH is reported to be 550 mV (ref. 22). Its axial ligand variant, M148L Rc, exhibits a potential of 613 mV versus a normal hydrogen electrode (NHE) at pH 6.2 (ref. 23), which may be directly compared to  $668 \pm 1$  mV at pH 6.0 for N47S/F114N/M121L Az. Therefore, N47S/F114N/M121L Az has a redox potential exceeding that of any known Rc variant at the same pH. Following the observed trend of increasing redox potential with decreasing pH, a maximum redox potential of  $706 \pm 3$  mV was observed for N47S/F114N/M121L Az at pH 4.0.

To investigate the combined effect of N47S and F114N on the structure of Az, a crystal structure of the N47S/F114N variant was obtained (Fig. 1c). The structure showed that the F114N mutation puts a hydrogen-bond donor near the backbone carbonyl of Gly 45. X-ray crystallography is, however, not able to distinguish between oxygen and nitrogen atoms in a molecule. If the positions of the oxygen and nitrogen on the side chain of Asn 114 are reversed, the amide nitrogen is in position to disrupt hydrogen-bonding interactions between the copper ligand, His 117, and other parts of the protein. Rc and other high potential cupredoxins achieve a similar effect by crowding the area around this His residue with Phe residues<sup>24</sup>.

Having achieved predictable tuning of the redox potential of Az to the high end of cupredoxin potentials, we also wanted to tune its potential to the low end. It is known that replacing the axial ligand Met 121 with Gln, the axial ligand found in low potential stellacyanin, lowers the potential of Az to  $\sim 190$  mV (refs 25, 26). Another report has shown that mutation of Phe 114 in Az to a Pro residue causes a reorganization of the hydrogen-bonding network near the copper binding site of Az, which lowers the redox potential by  $\sim 90$  mV. This decrease is due to the deletion of a direct hydrogen bond to Cys 112, which increases the electron density on the sulphur ligand, thereby lowering the redox potential<sup>27</sup>. Hence, we constructed F114P/M121Q and F114P/M121L Az, demonstrating that the effect of F114P is also additive with that of the axial ligand hydrophobicity (Fig. 2a). With a  $E^\circ$  of  $90 \pm 8$  mV at pH 7.0, F114P/M121Q Az boasts the lowest  $E^\circ$  reported for all cupredoxins and their variants. Following the same pH trend observed for the N47S/F114N/M121L Az variant, a minimum redox value of  $-2 \pm 13$  mV was observed at pH 9.0. We determined the crystal structure of F114P/M121Q Az (Fig. 1d), which shows similar effects to those seen in F114P Az: a disrupted hydrogen bond between Phe 114 and Cys 112.

In summary, we have predictably and rationally tuned the  $E^\circ$  of Az to the full range reported for native cupredoxins and beyond, as well as many  $E^\circ$  in between (Fig. 2b). This unprecedented level of control over an electron transfer protein was achieved by mapping out major interactions, an approach that may apply to all other redox proteins. While redox agents for use in organic solvents are plentiful, water-soluble and stable redox agents are rare, and those available display a limited potential range. Previously, to cover a wide potential range, different compounds with different surface interactions or electron transfer properties had to be used. As point mutants of a given protein have almost identical structures and surface properties, such proteins will find a wide range of applications as redox agents in areas such as fundamental biochemical and biophysical studies of enzymes, and could be tailor-made for use as artificial photosynthetic centres and fuel cell catalysts for energy conversion.

## METHODS SUMMARY

Preparation of all Az variants was performed using previously published protocols<sup>28,29</sup>. Reduction potentials were measured by absorbing copper incorporated protein onto graphite electrodes and scanning for redox events with cyclic

voltammetry versus a Ag/AgCl reference electrode. Observed reduction potentials were corrected to an NHE by adding 210 mV. Owing to pH dependent variation in the reduction potential of cupredoxin proteins, all reduction potentials were compared at the same pH unless otherwise noted. For a more complete comparison of absorption spectroscopy, the extinction coefficients of each absorbance were calculated by quantifying the amount of protein in each sample with inductively coupled plasma optical emission spectrometry. EPR spectra were obtained as a glass at 30 K. Nuclear magnetic resonance (NMR) spectra of the reduced proteins were collected in buffered D<sub>2</sub>O solution.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 3 July; accepted 30 September 2009.**

- Beratan, D. N., Onuchic, J. N., Winkler, J. R. & Gray, H. B. Electron-tunneling pathways in proteins. *Science* **258**, 1740–1741 (1992).
- Kanan, M. W. & Nocera, D. G. In situ formation of an oxygen-evolving catalyst in neutral water containing phosphate and Co<sup>2+</sup>. *Science* **321**, 1072–1075 (2008).
- Blanford, C. F., Heath, R. S. & Armstrong, F. A. A stable electrode for high-potential, electrocatalytic O<sub>2</sub> reduction based on rational attachment of a blue copper oxidase to a graphite surface. *Chem. Commun.* 1710–1712 (2007).
- Gray, H. B., Malmström, B. G. & Williams, R. J. P. Copper coordination in blue proteins. *J. Biol. Inorg. Chem.* **5**, 551–559 (2000).
- Solomon, E. I., Szilagyi, R. K., DeBeer George, S. & Basumallick, L. Electronic structures of metal sites in proteins and models: contributions to function in blue copper proteins. *Chem. Rev.* **104**, 419–458 (2004).
- Vila, A. J. & Fernández, C. O. in *Handbook on Metalloproteins* (eds Bertini, I., Sigel, A. & Sigel, H.) 813–856 (Marcel Dekker, 2001).
- Lu, Y. in *Comprehensive Coordination Chemistry II: From Biology to Nanotechnology* Vol. 8, *Biocoordination Chemistry* (eds Que, L. & Tolman, W. B.) 91–122 (Elsevier, 2004).
- Dennison, C. Investigating the structure and function of cupredoxins. *Coord. Chem. Rev.* **249**, 3025–3054 (2005).
- Solomon, E. I. Spectroscopic methods in bioinorganic chemistry: blue to green to red copper sites. *Inorg. Chem.* **45**, 8012–8025 (2006).
- Varadarajan, R., Zewert, T. E., Gray, H. B. & Boxer, S. G. Effects of buried ionizable amino acids on the reduction potential of recombinant myoglobin. *Science* **243**, 69–72 (1989).
- Battistuzzi, G., Borsari, M., Loschi, L., Righi, F. & Sola, M. Redox thermodynamics of blue copper proteins. *J. Am. Chem. Soc.* **121**, 501–506 (1999).
- Dey, A. et al. Solvent tuning of electrochemical potentials in the active sites of HiPIP versus ferredoxin. *Science* **318**, 1464–1468 (2007).
- Garner, D. K. et al. Reduction potential tuning of the blue copper center in *Pseudomonas aeruginosa* azurin by the axial methionine as probed by unnatural amino acids. *J. Am. Chem. Soc.* **128**, 15608–15617 (2006).
- Berry, S. M., Ralle, M., Low, D. W., Blackburn, N. J. & Lu, Y. Probing the role of axial methionine in the blue copper center of azurin with unnatural amino acids. *J. Am. Chem. Soc.* **125**, 8760–8768 (2003).
- Abdelhamid, R. F. et al.  $\pi$ - $\pi$  interaction between aromatic ring and copper-coordinated His81 imidazole regulates the blue copper active-site structure. *J. Biol. Inorg. Chem.* **12**, 165–173 (2007).
- Yanagisawa, S., Sato, K., Kikuchi, M., Kohzuma, T. & Dennison, C. Introduction of a  $\pi$ - $\pi$  interaction at the active site of a cupredoxin: characterization of the Met16Phe pseudoazurin mutant. *Biochemistry* **42**, 6853–6862 (2003).
- Kanbi, L. D. et al. Crystal structures of the Met148 and Ser86Asp mutants of rusticyanin from *Thiobacillus ferrooxidans*: insights into the structural relationship with the cupredoxins and the multi copper proteins. *J. Mol. Biol.* **320**, 263–275 (2002).
- Hall, J. F., Kanbi, L. D., Harvey, I., Murphy, L. M. & Hasnain, S. S. Modulating the redox potential and acid stability of rusticyanin by site-directed mutagenesis of Ser86. *Biochemistry* **37**, 11451–11458 (1998).
- Grossmann, J. G., Ingledew, W. J., Harvey, I., Strange, R. W. & Hasnain, S. S. X-ray absorption studies and homology modeling define the structural features that specify the nature of the copper site in rusticyanin. *Biochemistry* **34**, 8406–8414 (1995).
- Walter, R. L. et al. Multiple wavelength anomalous diffraction (MAD) crystal structure of rusticyanin: a highly oxidizing cupredoxin with extreme acid stability. *J. Mol. Biol.* **263**, 730–751 (1996).
- Lowery, M. D. & Solomon, E. I. Axial ligand bonding in blue copper proteins. *Inorg. Chim. Acta* **198–200**, 233–243 (1992).
- Alcaraz, L. A. et al. Folding and unfolding in the blue copper protein rusticyanin: role of the oxidation state. *Bioinorg. Chem. Appl.* **1–9** (2007).
- Hall, J. F., Kanbi, L. D., Strange, R. W. & Hasnain, S. S. Role of the axial ligand in type 1 Cu centers studied by point mutations of Met148 in rusticyanin. *Biochemistry* **38**, 12675–12680 (1999).
- Giudici-Ortoni, M. T., Guerlesquin, F., Bruschi, M. & Nitschke, W. Interaction-induced redox switch in the electron transfer complex rusticyanin-cytochrome *c*<sub>4</sub>. *J. Biol. Chem.* **274**, 30365–30369 (1999).
- Hart, P. J. et al. A missing link in cupredoxins: crystal structure of cucumber stellacyanin at 1.6 Å resolution. *Protein Sci.* **5**, 2175–2183 (1996).

26. Romero, A. *et al.* X-ray analysis and spectroscopic characterization of M121Q azurin. A copper site model for stellacyanin. *J. Mol. Biol.* **229**, 1007–1021 (1993).
27. Yanagisawa, S., Banfield, M. J. & Dennison, C. The role of hydrogen bonding at the active site of a cupredoxin: the Phe114Pro azurin variant. *Biochemistry* **45**, 8812–8822 (2006).
28. Chang, T. K. *et al.* Gene synthesis, expression, and mutagenesis of the blue copper proteins azurin and plastocyanin. *Proc. Natl Acad. Sci. USA* **88**, 1325–1329 (1991).
29. Mizoguchi, T. J., Di Bilio, A. J., Gray, H. B. & Richards, J. H. Blue to type 2 binding. Copper(II) and Cobalt(II) derivatives of a Cys112Asp mutant of *Pseudomonas aeruginosa* azurin. *J. Am. Chem. Soc.* **114**, 10076–10078 (1992).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** This material is based on work supported by the US National Science Foundation under award no. CHE 05-52008. We thank E. Marshall, D. Poe, A. Huang and J. Li for discussions, and for help in protein expression and

purification, and in spectroscopic and electrochemical data collection. N.M.M. is an NIH Predoctoral trainee, supported by the National Institutes of Health under Ruth L. Kirschstein National Research Service Award 5 T32 GM070421 from the National Institute of General Medical Sciences.

**Author Contributions** N.M.M. performed most of the experimentation, led the study and authored most of the manuscript. D.K.G. and T.D.W. contributed intellectually and assisted in experimentation and editing. Crystal diffraction patterns were collected by H.R. and refined into PDB structures by Y.-G.G. M.J.N. assisted with EPR data collection and simulation. Y.L. designed and guided the project, and edited the paper.

**Author Information** Crystallographic data have been deposited at the Protein Data Bank with the following PDB codes: N47S/M121L Az, 3JT2; N47S/F114N Az, 3JTB; F114P/M121Q Az, 3INO. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to Y.L. ([yi-lu@illinois.edu](mailto:yi-lu@illinois.edu)).



## METHODS

**Protein expression.** Az mutants were constructed, expressed and purified as apo-protein using a previously reported protocol<sup>28,29</sup>. Copper was added by slow titration of the apo-protein with 5 mM CuSO<sub>4</sub> and monitoring the increase in intensity of the visible peak around 620 nm. Full incorporation was considered to be the point at which adding more CuSO<sub>4</sub> resulted in no further increase in the intensity of the ~620 nm peak. The resulting copper incorporated proteins were purified from free copper ions through a gel filtration column (PD-10).

**Reduction potential measurements.** Reduction potentials were recorded using a pyrolytic graphite edge (PGE) electrode assembled according to literature procedures<sup>30</sup> with a CH Instruments model 620A Electrochemical Analyser potentiostat at a scan rate of 0.25 V s<sup>-1</sup> versus an Ag/AgCl reference. No dependence of the peak position on scan rate was observed for any of the proteins. Cyclic voltammetry of mutants containing both N47S and F114N, which were observed to be unstable on the PGE electrode, were collected on a less destabilizing glassy pyrolytic graphite electrode. After polishing of the electrodes with alumina on a nylon polishing pad and sonication to remove excess alumina, 2 µl of 2 mM protein solution were added to the graphite electrode. Alternatively, a 200 µM solution of the proteins containing both N47S and F114N were used due to the aforementioned lower stability of these proteins. Incubation of the protein on the electrode for 3 min at 4 °C was sufficient to absorb enough protein for a single cyclic voltammetry run. The electrodes were then immersed in buffer containing 50 mM NH<sub>4</sub>OAc, 40 mM MOPS, 40 mM MES, 40 mM Tris, 40 mM CAPS and 100 mM NaCl. Potentials were corrected to NHE by adding 210 mV to the observed potentials. Spectra were baseline corrected with the SOAS software package.

**Electronic absorption and EPR spectroscopy.** Ultraviolet–visible absorption spectra were obtained on a Varian Cary 5000 spectrometer at 10 °C and pH 6.0. Extinction coefficients were calculated by quantifying sulphur and copper from the purified protein via inductively coupled plasma optical emission spectrometry. X-band EPR spectra were collected at 30 K as a glass in 50% glycerol on a Varian 122 spectrometer with an Air Products Helitran cryostat. EPR g-values were calculated by simulating the spectra with SIMPOW<sup>31,32</sup>.

**NMR spectroscopy.** NMR samples of reduced azurin variants were prepared according to a previously described method in buffered D<sub>2</sub>O solution<sup>16</sup>. Owing to the much lowered redox potential of F114P/M121Q azurin, dithionite was used as a reductant rather than ascorbic acid. Spectra were obtained on a Varian Unity 500 spectrometer.

**Crystallization of Az variants.** Each protein was crystallized at pH 6.0 by hanging drop vapour diffusion. A 2 µl portion of 2 mM protein solution was mixed with 2 µl of well buffer solution consisting of 80 mM sodium acetate, 0.24 M calcium chloride, 0.24 M lithium nitrate. Various masses and concentrations of polyethylene glycol (PEG) were screened for crystal formation. The highest quality crystals formed in 25% PEG 8000 and PEG 4000.

30. Jeuken, L. J. C. & Armstrong, F. A. Electrochemical origin of hysteresis in the electron-transfer reactions of adsorbed proteins: contrasting behavior of the “blue” copper protein, azurin, adsorbed on pyrolytic graphite and modified gold electrodes. *J. Phys. Chem. B* **105**, 5271–5282 (2001).
31. Nilges, M. J. *Electron Paramagnetic Resonance Studies of Low Symmetry Nickel(I) and Molybdenum(V) Complexes*. Thesis, Univ. Illinois (1979).
32. Chang, H. R. et al. An unusually stable manganese(II)manganese(III) complex with novel EPR spectra: synthesis, structure, magnetism, and EPR analysis. *J. Am. Chem. Soc.* **110**, 625–627 (1988).

# Structural basis of inter-protein electron transfer for nitrite reduction in denitrification

Masaki Nojiri<sup>1</sup>, Hiroyasu Koteishi<sup>1</sup>, Takuya Nakagami<sup>1</sup>, Kazuo Kobayashi<sup>2</sup>, Tsuyoshi Inoue<sup>3</sup>, Kazuya Yamaguchi<sup>1</sup> & Shinnichiro Suzuki<sup>1</sup>

Recent earth science studies have pointed out that massive acceleration of the global nitrogen cycle by anthropogenic addition of bio-available nitrogen has led to a host of environmental problems<sup>1</sup>. Nitrous oxide (N<sub>2</sub>O) is a greenhouse gas that is an intermediate during the biological process known as denitrification<sup>2</sup>. Copper-containing nitrite reductase (CuNIR) is a key enzyme in the process; it produces a precursor for N<sub>2</sub>O by catalysing the one-electron reduction of nitrite (NO<sub>2</sub><sup>−</sup>) to nitric oxide (NO)<sup>3</sup>. The reduction step is performed by an efficient electron-transfer reaction with a redox-partner protein<sup>4–6</sup>. However, details of the mechanism during the electron-transfer reaction are still unknown. Here we show the high-resolution crystal structure of the electron-transfer complex for CuNIR with its cognate cytochrome *c* as the electron donor. The hydrophobic electron-transfer path is formed at the docking interface by desolvation owing to close contact between the two proteins. Structural analysis of the interface highlights an essential role for the loop region with a hydrophobic patch for protein–protein recognition; it also shows how interface construction allows the variation in atomic components to achieve diverse biological electron transfers.

The biological reduction of nitrite ions to gaseous nitrogen monoxide (NO<sub>2</sub><sup>−</sup> + 2H<sup>+</sup> + e<sup>−</sup> → NO + H<sub>2</sub>O) is catalysed by two very different enzymes called haem *cd*<sub>1</sub>-containing nitrite reductase (*cd*<sub>1</sub>NIR) and CuNIR, which are less commonly known by their product gene names, NirS and NirK, respectively<sup>3</sup>. They are mutually exclusive at the species level in the denitrifying bacteria distributed widely on Earth<sup>7</sup>. For the one-electron reduction of NO<sub>2</sub><sup>−</sup>, the enzyme is linked to respiratory electron-transfer chains by cytochrome *c* (Cyt *c*) and/or cupredoxin in the cell<sup>8,9</sup>. Although the electrons destined for the enzyme pass through the cytochrome *bc*<sub>1</sub> complex in the chain, how they pass on from this complex to the enzyme is still a matter of uncertainty. In general, CuNIR folds a homotrimeric structure with two distinct Cu-binding sites through an approximately 37-kilodalton (kDa) monomer unit<sup>10</sup>. The type 1 Cu site (T1Cu) buried within each monomer relays an electron from the redox-partner protein to the catalytic type 2 Cu site (T2Cu), where NO<sub>2</sub><sup>−</sup> is reduced to NO. This enzyme is further classified into two subgroups based on the spectroscopic properties of the T1Cu site, which are blue and green<sup>11</sup>. Whereas the physiological redox-partner protein for green CuNIR is pseudoazurin (PAz), the blue CuNIR has been suggested to accept an electron from azurin (Az) or Cyt *c*<sup>11–13</sup>.

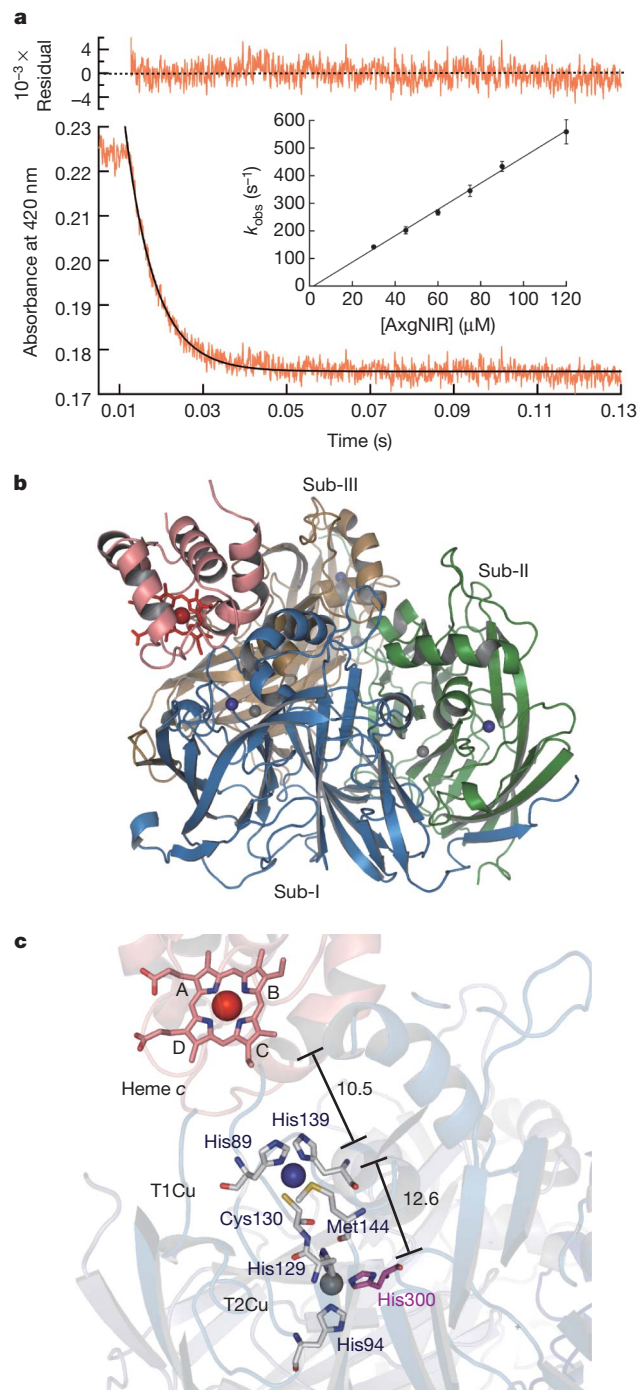
In a ubiquitous Gram-negative non-fermenting rod, *Achromobacter xylosoxidans* GIFU 1051 (the organism formerly known as *Alcaligenes xylosoxidans*)<sup>14</sup>, it has been found that two structural genes (*nirK* and *nirM*) of the blue CuNIR (AxxNIR) and Cyt *c*<sub>551</sub> constitute an operon presumed to be a redox-partner pair

(Supplementary Fig. 1). The *nirM* gene is also part of the operon that harbours the *nirS* gene coding for *cd*<sub>1</sub>NIR in other denitrifying bacteria, and the product Cyt *c*<sub>551</sub> protein acts as a physiological redox partner for *cd*<sub>1</sub>NIR<sup>15</sup>. Recent gene disruption studies on the *nirK* gene-harbouring bacterium *Bradyrhizobium japonicum* have shown that Cyt *c*<sub>550</sub> is required for growth under denitrifying conditions<sup>16</sup>. Moreover, earlier work by Zumft and colleagues has demonstrated that CuNIR can function instead of *cd*<sub>1</sub>NIR; this is most likely caused by the use of Cyt *c* as a redox partner in *Pseudomonas stutzeri*<sup>17</sup>. Thus, it is reasonable to assume that not only Az but also Cyt *c*<sub>551</sub> plays a role as an *in vivo* redox partner for AxxNIR.

To elucidate the physiological relationship between AxxNIR and Cyt *c*<sub>551</sub>, the formation of a functional electron-transfer complex between these proteins was analysed by stopped-flow kinetics and a crystallographic approach (Fig. 1). The kinetics of electron transfer from the reduced Cyt *c*<sub>551</sub> to AxxNIR was monitored at a wavelength of 420 nm originating from a Soret band of the haem *c* group under anaerobic conditions. The rapid decay of the absorbance is due to the oxidation of Cyt *c*<sub>551</sub> and concomitant reduction of AxxNIR (Fig. 1a). The second-order electron-transfer rate constant between the two proteins,  $(4.8 \pm 0.2) \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$  (s.d.) (Fig. 1a, inset), strongly supports the idea that Cyt *c*<sub>551</sub> interacts functionally with AxxNIR as an electron donor. To understand the detailed mechanism further, the crystal structure of the binary complex of AxxNIR with Cyt *c*<sub>551</sub> was determined at a resolution of 1.7 Å (Fig. 1b). A model of the complex, which consists of one Cyt *c*<sub>551</sub> and one trimeric AxxNIR molecule within an asymmetric unit, contains 8,242 protein atoms, one haem *c* group, and six copper atoms (Fig. 1b; see also Supplementary Figs 2 and 3a). The AxxNIR molecules contribute to major contacts in the crystal lattice and provide only one space for Cyt *c*<sub>551</sub>; this is bound to its docking site on a subunit (Sub-I) of AxxNIR without steric hindrance (Supplementary Fig. 3b).

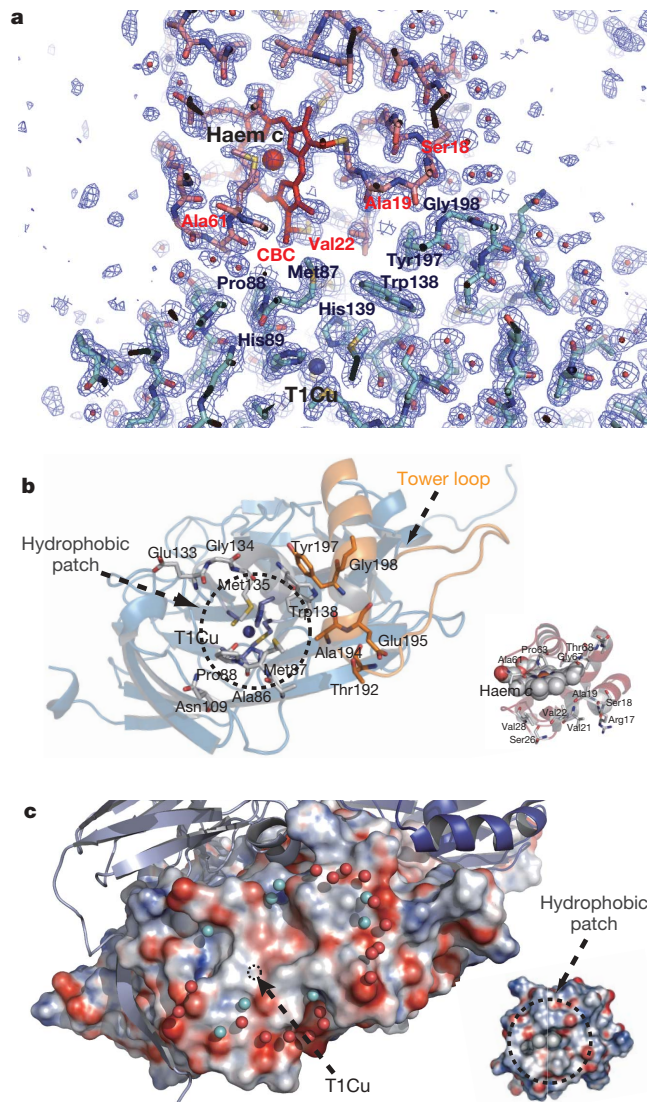
Within the docking interface, the change in solvent-accessible surface area upon complex formation is estimated to be  $-529 \text{ Å}^2$  for Sub-I and  $-553 \text{ Å}^2$  for Cyt *c*<sub>551</sub>; the gap index (the ratio of gap volume to buried interface solvent-accessible surface area) is 3.55 Å. The haem *c* group partly protrudes from the protein interior of the Cyt *c*<sub>551</sub> molecule at two parts: the thioether-bonded substituent linked to pyrrole ring C and the propionate groups on pyrrole rings A and D (Fig. 1c). The Cyt *c*<sub>551</sub>-docking site near the T1Cu site of Sub-I and the haem *c* group are in close contact at a 3.5-Å distance between the Cε atom of Met 87 in Sub-I and the edge CBC methyl carbon of haem *c* (Fig. 2a). At least ten amino-acid residues of AxxNIR are associated with at least 11 residues and the haem *c* group of Cyt *c*<sub>551</sub> at the interface (Supplementary Table 2). There is no salt bridge at the interface, and only three direct hydrogen bonds were observed. In general, direct hydrogen bonds between the docking

<sup>1</sup>Department of Chemistry, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan. <sup>2</sup>The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567-0047, Japan. <sup>3</sup>Department of Materials Chemistry, Graduate School of Engineering, Osaka University, Suita, Osaka 565-0871, Japan.



**Figure 1 | The electron-transfer complex between AxxNIR and Cyt  $c_{551}$ .** **a**, Stopped-flow kinetics of inter-protein electron-transfer reaction from reduced Cyt  $c_{551}$  to AxxNIR. The orange line represents the observed data; the black line is a fitting curve. The inset shows plots of  $k_{\text{obs}}$  versus AxxNIR concentrations. Error bars denote s.d. **b**, Overall structure of the binary complex of AxxNIR with Cyt  $c_{551}$ . The Cyt  $c_{551}$  molecule is represented as a pink-coloured ribbon, the Cyt  $c_{551}$ -docked Sub-I of AxxNIR is shown in sky blue, the undocked Sub-II in green and the Sub-III in dark-yellow. The haem group (red), T1Cu (dark blue) and T2Cu (grey) are depicted as balls and sticks. **c**, Arrangement of three redox centres. The haem group and ligand residues of both Cu centres are shown as sticks. The distances from the CBC methyl carbon of haem  $c$  to T1Cu and from T1Cu to T2Cu are given in ångströms.

proteins are unfavourable for a transient electron-transfer complex because of energetically disadvantageous desolvation<sup>18</sup>; namely, the interface of the reaction-centre:Cyt  $c_2$  complex also has three direct

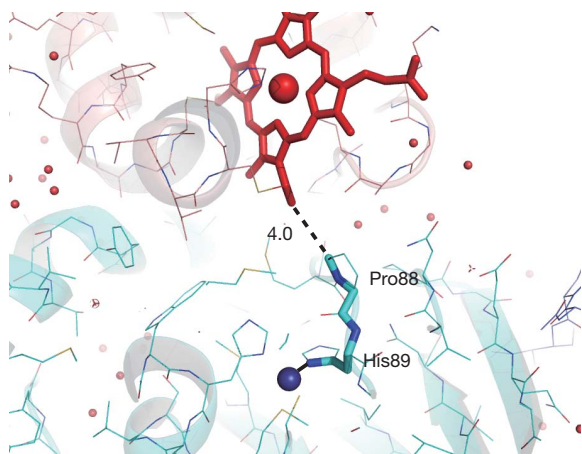


**Figure 2 | The docking interface.** **a**, Electron-density map of the docking interface between AxxNIR and Cyt  $c_{551}$ . A  $\sigma_A$ -weighted  $2F_{\text{obs}} - F_{\text{calc}}$  map contoured at  $1.0\sigma$  is shown around the interface. **b**, The building block of the interface. The amino-acid residues at the interface and the T1Cu ligands are represented as sticks. The T1Cu ligands, the hydrophobic patch and the 'tower loop' are coloured in blue, grey and orange, respectively. Inset: the residues around the haem group in Cyt  $c_{551}$ . The haem group is shown as spheres. **c**, Electrostatic potentials of contact surfaces. Eight water molecules bridging between both proteins through hydrogen bonds are coloured in cyan and the other waters binding to Sub-I or Cyt  $c_{551}$  in red. Electrostatic potential was calculated by the program APBS plugged-in to PyMOL. The displayed potentials range from  $-15$  (red) to  $15$  (blue)  $kT e^{-1}$ .

hydrogen bonds<sup>19</sup> and that of the yeast Cyt  $c$  peroxidase (CcP):Cyt  $c$  complex one<sup>20</sup>, but that of the Cyt  $bc_1$ :Cyt  $c$  complex has no hydrogen bond<sup>21</sup>.

The interacting residues of AxxNIR are localized both at the hydrophobic patch and at the 'tower loop' region<sup>22</sup> near the T1Cu site (Fig. 2b). Contact between the hydrophobic patches in both proteins brings their redox centres within  $10.5$  Å, which is close enough to allow rapid electron transfer<sup>23</sup>. Eight water molecules bridging the two proteins through hydrogen bonds at the interface stabilize the protein-protein interaction. All of the water molecules at the interface form a characteristic semicircle around the hydrophobic patch, and the non-polar core interface is sealed off from the aqueous environment (Fig. 2c). A comparison of the contact residues of Cyt



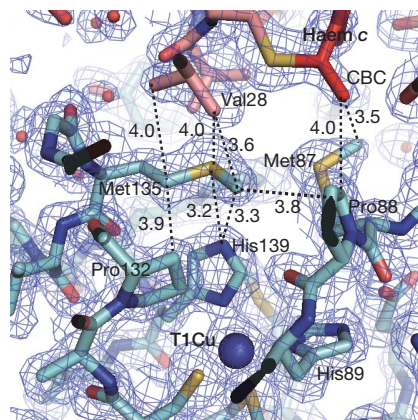


**Figure 3 | A dominant electron-transfer path in the AxxNIR:Cyt  $c_{551}$  complex.** The best path predicted by PATHWAY<sup>24</sup> is shown as a broken line (through-space process) and sticks (through-bond process). The distance of the through-space jump between the CBC methyl group and the C $\delta$  atom of Pro 88 is given in ångströms. The haem group (red), T1Cu atom (dark blue) and water molecules (red) are depicted as balls and sticks.

$c_{551}$  with those of Cyt  $c$  in the previously known electron-transfer complexes<sup>19–21</sup> reveals that the van der Waals interaction using the short side-chain residues (Ala, Thr and Val) around the haem CBC methyl group is one of the consensus features of the Cyt  $c$  molecules; these molecules need a close interaction between the redox partners in their electron-transfer reactions.

PATHWAY analysis<sup>24</sup> illustrates a dominant electron-transfer path through the entry/exit port inside the hydrophobic core at the interface (Fig. 3); the associated theoretical electron-transfer parameters are listed in Supplementary Table 3. An electron that leaves iron by the exposed CBC methyl carbon is directly transferred to the C $\delta$  atom of Pro 88 of AxxNIR by a through-space jump; it then shifts from Pro 88 to T1Cu through the His 89 ligand. The *Alcaligenes faecalis* CuNIR(AfNIR):PAz complex structure proposed by a recent NMR study<sup>25</sup> reveals that a similar hydrophobic core surrounded by polar interactions is formed at the interface as well as the AxxNIR:Cyt  $c_{551}$  complex. However, the predicted electron-transfer paths in the AfNIR:PAz complex seem to be slightly different from that in the AxxNIR:Cyt  $c_{551}$  complex (Supplementary Fig. 4). This disagreement is probably due to the directional and structural differences between the redox centres of PAz and Cyt  $c_{551}$ . For instance, the haem orientation that determines the position of the CBC methyl carbon (electron exit port) is a key factor in selecting the target electron-entry port atom of AxxNIR for an inter-protein through-space electron-transfer jump in the AxxNIR:Cyt  $c_{551}$  complex. Moreover, the concept of ‘pseudospecificity’ has often been invoked to explain how several donor proteins with very different structures such as PAz and Cyt  $c$  can transfer electrons to the same enzyme ( $cd_1$ NIR)<sup>26</sup>.

Pseudospecificity suggests that only a transient and promiscuous interaction between the enzyme surface patches and its partner proteins is required for inter-protein electron transfer to occur. However, flexible partner-recognition inducing some slight structural perturbations observed in the AxxNIR:Cyt  $c_{551}$  structure (Supplementary Fig. 5) is likely to keep the balance of specific and fast association/dissociation between both proteins and affect the nature of the coupling between each specific patch of the proteins for the inter-protein electron-transfer reaction<sup>23,24</sup>. For example, the Met 135 residue in the Cyt  $c_{551}$ -docked Sub-I has a side-chain structure tilted towards the His 139 ligand of T1Cu (Fig. 4). The S $\delta$  atom of Met 135 is within 3.2 Å of the Ne2 atom of His 139 and is tightly fixed by covering with Val 28 of Cyt  $c_{551}$ . Consequently, His 139 is no longer exposed to the solvent in the complex. Although the side chain of Met 135 shows multiple conformations in the partner-free state (Supplementary Fig. 6), it is locked at the interface of the complex.



**Figure 4 | Packing of Met 135 at the interface by Cyt  $c_{551}$  docking.** The Cyt  $c_{551}$ -docked Sub-I, Cyt  $c_{551}$ , and haem group are represented as cyan, pink and red sticks, respectively. T1Cu is shown as a dark-blue sphere. A  $\sigma_A$ -weighted  $2F_{\text{obs}} - F_{\text{calc}}$  map contoured at  $1.0\sigma$  is shown around the Met 135 residue intervened between two redox centres, haem and T1Cu. The distances of the van der Waals contacts (dotted line) around Met 135 are given in ångströms.

This finding seems to be a consensus feature for the AxxNIR:Cyt  $c_{551}$  and AfNIR:PAz complexes (Supplementary Fig. 4). The stopped-flow electron-transfer kinetics of the Met 135 to Ser mutant, which has a more compact side chain ( $-\text{CH}_2-\text{OH}$ ) for Ser than that ( $-\text{CH}_2-\text{CH}_2-\text{S}-\text{CH}_3$ ) for Met, indicates that the second-order electron-transfer rate constant of the mutant is less than that of native AxxNIR by a factor of about 8; this suggests that the atomic packing is very important for the haem-to-T1Cu electron-transfer reaction (Supplementary Fig. 7). The redox potentials of the haem  $c$  group of Cyt  $c_{551}$  and the T1Cu site of AxxNIR were independently determined to be +241 and +290 mV compared with a normal hydrogen electrode at pH 6.0, respectively<sup>11</sup>. Because it has been reported that the redox potential of T1Cu is considerably changed by perturbations of the His ligand<sup>27</sup>, the docking of partner protein onto CuNIR may tune the redox potential of T1Cu through structural constraint and desolvation of the His ligand for smooth electron transfer. This is observed in the redox potential tuning of the T1Cu sites of the methylamine dehydrogenase (MADH):amicyanin (Am) and MADH:Am:Cyt  $c_{551}$  complexes<sup>28</sup>.

Moreover, it has recently been found by genomic analysis that some CuNIRs are covalently attached to a Cyt  $c$  domain at the carboxy (C) terminus<sup>29</sup>. Viewing the CuNIR-Cyt  $c$  fusion event from the perspective of the *nirK-nirM* operon allows us to speculate that CuNIR may have captured the Cyt  $c_{551}$  gene for construction of an effective self-electron-transfer system, underlying the fact that the fusion of two genes in one organism is directly correlated with the functional association of the same gene products in another<sup>30</sup>. The AxxNIR:Cyt  $c_{551}$  complex structure is expected to serve as an archetypal template for understanding how intra-protein haem-to-T1Cu electron transfer occurs in the CuNIR-Cyt  $c$  fusion protein.

## METHODS SUMMARY

AxxNIR and Cyt  $c_{551}$  were directly purified from the cell-free extract of the growth cells for *A. xylosoxidans* GIFU 1051. Stopped-flow kinetics of the electron-transfer reaction from reduced Cyt  $c_{551}$  to AxxNIR was performed in a 10 mM potassium phosphate buffer (pH 6.0) at 25 °C under anaerobic conditions. To keep the pseudo-first-order conditions, the concentration ratios of AxxNIR to Cyt  $c_{551}$  were always more than ten times (AxxNIR, 30–120  $\mu\text{M}$ ; Cyt  $c_{551}$ , 1.0  $\mu\text{M}$ ). The pseudo-first-order rate constants ( $k_{\text{obs}}$ ) were obtained by fitting the experimental data with a single exponential function. The second-order electron-transfer rate constant was estimated from the slope of the plots of AxxNIR concentrations versus  $k_{\text{obs}}$ . Crystals of the binary AxxNIR:Cyt  $c_{551}$  complex grew in space group  $P2_12_12_1$  with  $a = 63.4$  Å,  $b = 103.9$  Å and  $c = 163.1$  Å; they contained one AxxNIR trimer molecule and one Cyt  $c_{551}$  molecule in an asymmetric unit. The final  $R$  and  $R$ -free factors are 16.5% and 19.5% at a resolution of 1.7 Å, respectively. In the final model, 90.2% of the residues are

in the most favourable regions of the Ramachandran plot, with no residues in disallowed regions (Supplementary Table 1).

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Received 30 April; accepted 14 September 2009.**

- Gruber, N. & Galloway, J. N. An earth-system perspective of the global nitrogen cycle. *Nature* **451**, 293–296 (2008).
- Firestone, M. K., Firestone, R. B. & Tiedje, J. M. Nitrous oxide from soil denitrification: factors controlling its biological production. *Science* **208**, 749–751 (1980).
- Zumft, W. G. Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev.* **61**, 533–616 (1997).
- Kataoka, K. *et al.* Structure-based engineering of *Alcaligenes xylosoxidans* copper-containing nitrite reductase enhances intermolecular electron transfer reaction with pseudoazurin. *J. Biol. Chem.* **279**, 53374–53378 (2004).
- Murphy, L. M., Dodd, F. E., Yousafzai, F. K., Eady, R. R. & Hasnain, S. S. Electron donation between copper containing nitrite reductases and cupredoxins: the nature of protein–protein interaction in complex formation. *J. Mol. Biol.* **315**, 859–871 (2002).
- Kukimoto, M. *et al.* Site-directed mutagenesis of azurin from *Pseudomonas aeruginosa* enhances the formation of an electron-transfer complex with a copper-containing nitrite reductase from *Alcaligenes faecalis* S-6. *FEBS Lett.* **394**, 87–90 (1996).
- Coyne, M. S., Arunakumari, A., Averill, B. A. & Tiedje, J. M. Immunological identification and distribution of dissimilatory heme *cd*<sub>1</sub> and nonheme copper nitrite reductases in denitrifying bacteria. *Appl. Environ. Microbiol.* **55**, 2924–2931 (1989).
- Vijsboom, E., Busch, J. E. & Canters, G. W. *In vivo* studies disprove an obligatory role of azurin in denitrification in *Pseudomonas aeruginosa* and show that *azu* expression is under control of RpoS and ANR. *Microbiology* **143**, 2853–2863 (1997).
- Ferguson, S. J. & Richardson, D. J. in *Respiration in Archaea and Bacteria*, Vol. 2 (ed. Zannoni, D.) 169–206 (Springer, 2004).
- Godden, J. W. *et al.* The 2.3 Å X-ray structure of nitrite reductase from *Achromobacter cycloclastes*. *Science* **253**, 438–442 (1991).
- Suzuki, S., Kataoka, K. & Yamaguchi, K. Metal coordination and mechanism of multi-copper nitrite reductase. *Acc. Chem. Res.* **33**, 728–735 (2000).
- Dodd, F. E. *et al.* Evidence for two distinct azurins in *Alcaligenes xylosoxidans* (NCIMB 11015): potential electron donors to nitrite reductase. *Biochemistry* **34**, 10180–10186 (1995).
- Deligeer, Kataoka K., Yamaguchi, K. & Suzuki, S. Spectroscopic and electrochemical properties of cytochrome *c*<sub>551</sub> from *Alcaligenes xylosoxidans* GIFU 1051. *Bull. Chem. Soc. Jpn* **73**, 1839–1840 (2000).
- Yabuuchi, E., Kawamura, Y., Kosako, Y. & Ezaki, T. Emendation of genus *Achromobacter* and *Achromobacter xylosoxidans* (Yabuuchi and Yano) and proposal of *Achromobacter ruhlandii* (Packer and Vishniac) comb. nov., *Achromobacter piechaudii* (Kiredjian *et al.*) comb. nov., and *Achromobacter xylosoxidans* subsp. *denitrificans* (Rüger and Tan) comb. nov. *Microbiol. Immunol.* **42**, 429–438 (1998).
- Hasegawa, N., Arai, H. & Igarashi, Y. Two *c*-type cytochromes, NirM and NirC, encoded in the *nir* gene cluster of *Pseudomonas aeruginosa* act as electron donors for nitrite reductase. *Biochem. Biophys. Res. Commun.* **288**, 1223–1230 (2001).
- Bueno, E., Bedmar, E. J., Richardson, D. J. & Delgado, M. J. Role of *Bradyrhizobium japonicum* cytochrome *c*<sub>550</sub> in nitrite and nitrate respiration. *FEMS Microbiol. Lett.* **279**, 188–194 (2008).
- Glockner, A. B., Jüngst, A. & Zumft, W. G. Copper-containing nitrite reductase from *Pseudomonas aureofaciens* is functional in a mutationally cytochrome *cd*<sub>1</sub>-free background (NirS<sup>−</sup>) of *Pseudomonas stutzeri*. *Arch. Microbiol.* **160**, 18–26 (1993).
- Miyashita, O., Onuchic, J. N. & Okamura, M. Y. Continuum electrostatic model for the binding of cytochrome *c*<sub>2</sub> to the photosynthetic reaction center from *Rhodospirillum rubrum*. *Biochemistry* **42**, 11651–11660 (2003).
- Axelrod, H. L. *et al.* X-ray structure determination of the cytochrome *c*<sub>2</sub>: reaction center electron transfer complex from *Rhodospirillum rubrum*. *J. Mol. Biol.* **319**, 501–515 (2002).
- Pelletier, H. & Kraut, J. Crystal structure of a complex between electron transfer partners, cytochrome *c* peroxidase and cytochrome *c*. *Science* **258**, 1748–1755 (1992).
- Solmaz, S. R. N. & Hunte, C. Structure of complex III with bound cytochrome *c* in reduced state and definition of a minimal core interface for electron transfer. *J. Biol. Chem.* **283**, 17542–17549 (2008).
- Boulanger, M. J. & Murphy, M. E. P. Crystal structure of the soluble domain of the major anaerobically induced outer membrane protein (AniA) from pathogenic *Neisseria*: a new class of copper-containing nitrite reductases. *J. Mol. Biol.* **315**, 1111–1127 (2002).
- Page, C. C., Moser, C. C., Chen, X. & Dutton, P. L. Natural engineering principles of electron tunnelling in biological oxidation-reduction. *Nature* **402**, 47–52 (1999).
- Onuchic, J. N., Beratan, D. N., Winkler, J. R. & Gray, H. B. Pathway analysis of protein electron-transfer reactions. *Annu. Rev. Biophys. Biomol. Struct.* **21**, 349–377 (1992).
- Vlasie, M. D., Fernández-Busnadiego, R., Prudêncio, M. & Ubbink, M. Conformation of pseudoazurin in the 152 kDa electron transfer complex with nitrite reductase determined by paramagnetic NMR. *J. Mol. Biol.* **375**, 1405–1415 (2008).
- Williams, P. A. *et al.* Pseudospecific docking surfaces on electron transfer proteins as illustrated by pseudoazurin, cytochrome *c*<sub>550</sub> and cytochrome *cd*<sub>1</sub> nitrite reductase. *Nature Struct. Biol.* **2**, 975–982 (1995).
- Canters, G. W. *et al.* The effect of pH and ligand exchange on the redox properties of blue copper proteins. *Faraday Discuss.* **116**, 205–220 (2000).
- Zhu, Z. *et al.* Molecular basis for interprotein complex-dependent effects on the redox properties of amicyanin. *Biochemistry* **37**, 17128–17136 (1998).
- Bertini, I., Cavallaro, G. & Rosato, A. Cytochrome *c*: occurrence and functions. *Chem. Rev.* **106**, 90–115 (2006).
- Enright, A. J., Iliopoulos, I., Kyripides, N. C. & Ouzounis, C. A. Protein interaction maps for complete genomes based on gene fusion events. *Nature* **402**, 86–90 (1999).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank A. Nakagawa, M. Suzuki, M. Yoshimura and E. Yamashita (beamline 44XU at Spring-8) for their support in the collection of X-ray data (proposal numbers 2007A6918 and 2007B6918, to M.N.). This work was supported in part by Grants-in-Aids for Scientific Research 20350078 (to S.S.) and Encouragement of Young Scientists 20750137 (to M.N.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and a Grant for Basic Science Research Projects from the Sumitomo Foundation (to M.N.).

**Author Contributions** M.N., H.K. and S.S. conceived and designed the project; M.N., H.K., T.N. and K.Y. purified; M.N. and H.K. crystallized; M.N., H.K., K.K. and T.I. conducted experimental works including data collection and structure analysis; M.N. and H.K. performed stopped-flow kinetics and analysed; M.N. and S.S. wrote the manuscript. All authors discussed the results and commented on the manuscript.

**Author Information** Atomic coordinates and structure factors for the reported crystal structure are deposited in the Protein Data Bank under accession number 2ZON. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to M.N. (nojiri@ch.wani.osaka-u.ac.jp) or S.S. (bic@ch.wani.osaka-u.ac.jp).

## METHODS

**Protein purification.** Purification of native Cu-containing nitrite reductase (AxxNIR) and cytochrome *c* (Cyt  $c_{551}$ ) from *A. xylosoxidans* GIFU 1051 was performed according to the procedures described previously<sup>31–34</sup>.

**Crystallization, data collection and processing.** Crystallization of the binary AxxNIR:Cyt  $c_{551}$  complex was performed at 4 °C using the hanging-drop vapour-diffusion method. The drops were prepared by mixing 0.5  $\mu$ l of AxxNIR (10 mg ml<sup>−1</sup> in 10 mM potassium phosphate, pH 6.0) and 0.5  $\mu$ l of Cyt  $c_{551}$  (10 mg ml<sup>−1</sup> in the same buffer) with 1  $\mu$ l of a reservoir solution containing 175 mM magnesium acetate tetrahydrate and 19% (w/v) PEG3350. The sample dishes were stored for crystallization in an incubator from 1 week to 1 month at 4 °C. Crystals of the AxxNIR:Cyt  $c_{551}$  complex were cryoprotected by transferring them to a reservoir solution supplemented with 10% (v/v) 2-methyl-2,4-pentandiol; they were then flash-frozen in liquid nitrogen for data collection at 100 K.

Diffraction data were collected at 100 K with a DIP-6040 detector (Bruker) on beamline 44XU at SPring-8 (Japan Synchrotron Radiation Research Institute). Image data were processed and scaled with the HKL package<sup>35</sup>. Crystals of the AxxNIR:Cyt  $c_{551}$  complex belong to the space group  $P2_12_12_1$  with the following cell parameters:  $a = 63.4$  Å,  $b = 103.9$  Å,  $c = 163.1$  Å. The data set was 99% complete overall to a resolution of 1.70 Å with an  $R_{\text{merge}}$  of 7.5%. The data processing statistics are summarized in Supplementary Table 1.

**Structure determination and refinement.** Initially, the structure of AxxNIR was determined through molecular replacement. The molecular structure of CuNIR (AxxNIR) from *A. xylosoxidans* NCIB11015, determined at a resolution of 1.04 Å (Protein Data Bank accession number 1oe1)<sup>36</sup>, was used as the search model. MOLREP<sup>37</sup> gave a unique molecular replacement solution, which was rigid-body refined using CNS<sup>38</sup>. One trimeric AxxNIR molecule was ultimately found in the asymmetric unit, and the resulting initial model gave an  $R_{\text{cryst}}$  of 28% at a resolution of 3.0 Å after rigid-body refinement and minimization. A  $F_{\text{obs}} - F_{\text{calc}}$  difference Fourier electron-density map calculated with the phase of the initial AxxNIR model gave strong signals sufficient to confirm the presence of a Cyt  $c_{551}$  molecule at the contour level of 2.0 $\sigma$  (Supplementary Fig. 2). The  $2F_{\text{obs}} - F_{\text{calc}}$  and  $F_{\text{obs}} - F_{\text{calc}}$  electron-density maps were used for building the Cyt  $c_{551}$  structure model into the densities at contour levels of 1.0 $\sigma$  ( $2F_{\text{obs}} - F_{\text{calc}}$ ) and 2.0 $\sigma$  ( $F_{\text{obs}} - F_{\text{calc}}$ ) using XtalView<sup>39</sup>. Several stages of model building and refinement were performed to trace the entire Cyt  $c_{551}$ , and the model was then thoroughly examined for possible errors using both maps at contour levels of 1.5 $\sigma$  ( $2F_{\text{obs}} - F_{\text{calc}}$ ) and  $\pm 3.0\sigma$  ( $F_{\text{obs}} - F_{\text{calc}}$ ). At a later stage, REFMAC<sup>40</sup> was used for refinement, and solvent molecules were gradually included in the model. Only solvent molecules with thermal parameters of  $< 50$  Å<sup>2</sup> and reasonable hydrogen-bonding properties were included in the final model. The structure was refined to obtain  $R_{\text{cryst}}$  and  $R_{\text{free}}$  of 0.16 and 0.19, respectively; root mean squared deviations of bond lengths and angles from the ideal values were 0.011 Å and 1.4°, respectively. The refined AxxNIR:Cyt  $c_{551}$  complex model was assessed using PROCHECK<sup>41</sup> and MOLPROBITY<sup>42</sup>. The average *B*-factors for all atoms of AxxNIR, Cyt  $c_{551}$  and water molecules were 16.5, 23.8 and 29.8 Å<sup>2</sup>, respectively. The amino (N)- and C-terminal residues of the model showed slightly elevated *B*-factors and somewhat diffuse electron density. In the final model, one residue (Gln) at the N terminus of AxxNIR, one residue (Arg) at the C terminus of AxxNIR and six residues (Ala-Ala-Asp-Ala-Pro-Ala) at the N terminus of Cyt  $c_{551}$  were undefined owing to disorder. The refinement statistics are summarized in Supplementary Table 1.

**Stopped-flow kinetics of electron-transfer reaction between AxxNIR and Cyt  $c_{551}$ .** The kinetic traces were acquired with an RA-2000 stopped-flow spectrophotometer (Otsuka Electronics) using the single-wavelength mode of the machine. Under a nitrogen atmosphere, 1.0  $\mu$ M of reduced Cyt  $c_{551}$  in 10 mM potassium phosphate buffer (pH 6.0) was loaded into one syringe of the stopped-flow apparatus and 30–120  $\mu$ M AxxNIR in the same buffer into the other. When the pressure was set to 5.0 kg cm<sup>−2</sup>, the initial 70 ms was required for delivery and complete mixing of the reactants. Pseudo-first-order rate constants ( $k_{\text{obs}}$ ) were calculated with a single exponential function using IgorPro version 4.02 (WaveMetrics).

**Mutation of Met 135 to Ser.** Site-directed mutagenesis was performed using the QuickChange site-directed mutagenesis kit (Stratagene). For a single mutant of Met 135 to Ser (M135S), the used oligonucleotide primers are 5'-CCGAAGGC AGCGTGCCCTGGCACGTGGTGTGCG-3' and 5'-CAGGGCAGCTGCCTTC GGGCGCGCAGTGGTA-3'. The plasmid pMaxNIR<sup>43</sup> was used as a template DNA. The integrity of the gene and the presence of the desired point mutation in the resulting plasmid pMaxNIR-M135S were confirmed by DNA sequencing.

The M135S mutant protein was expressed in *Escherichia coli* JM109 as a recombinant fusion protein with the maltose-binding protein; it was purified and digested with Factor-Xa to exclude maltose-binding protein, as described previously<sup>43</sup>. The electronic absorption spectrum was measured at room temperature with a Shimadzu UV-2450 spectrophotometer. Stopped-flow kinetics of the electron-transfer reaction between Cyt  $c_{551}$  and M135S were performed under identical conditions to the case of native AxxNIR.

- Iwasaki, H. *et al.* Cytochrome *c'* isolated from *Achromobacter xylosoxidans* GIFU 1055. *Plant Cell Physiol.* **27**, 733–736 (1986).
- Masuko, M., Iwasaki, H., Sakurai, T., Suzuki, S. & Nakahara, A. Characterization of nitrite reductase from a denitrifier, *Alcaligenes* sp. NCIB 11015. A novel copper protein. *J. Biochem.* **96**, 447–454 (1984).
- Suzuki, H. & Iwasaki, H. Studies on denitrification VI. Preparations and properties of crystalline blue protein and cryptocytocrome *c*, and role of copper in denitrifying enzyme from a denitrifying bacterium. *J. Biochem.* **52**, 193–199 (1962).
- Suzuki, S. *et al.* Spectroscopic characterization and intramolecular electron transfer processes of native and type 2 Cu-depleted nitrite reductases. *J. Biol. Inorg. Chem.* **2**, 265–274 (1997).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Ellis, M. J., Dodd, F. E., Sawers, G., Eady, R. R. & Hasnain, S. S. Atomic resolution structures of native copper nitrite reductase from *Alcaligenes xylosoxidans* and the active site mutant Asp92Glu. *J. Mol. Biol.* **328**, 429–438 (2003).
- Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022–1025 (1997).
- Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
- McRee, D. E. XtalView/Xfit—A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165 (1999).
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
- Lovell, S. C. *et al.* Structure validation by  $C\alpha$  geometry:  $\phi$ ,  $\psi$  and  $C\beta$  deviation. *Proteins* **50**, 437–450 (2003).
- Kataoka, K., Furusawa, H., Takagi, K., Yamaguchi, K. & Suzuki, S. Functional analysis of conserved aspartate and histidine residues located around the type 2 copper site of copper-containing nitrite reductase. *J. Biochem.* **127**, 345–350 (2000).



## PROSPECTS

# Career resilience

**It's not enough to be an expert on a specific topic. Today's scientists also need to be able to apply their knowledge, argues Peter Fiske.**

Some might assume that cranking out good results is all that is necessary to further a career. But such a 'career protocol' is not safe. That has never been more true than now, as the world tries to climb out of the current recession.

Although the economy has shown signs of improvement, scientists and engineers of all ages will be feeling the effects of the recession for some time. In the past, the scientific community was often insulated from the brunt of economic downturns. Buoyed by the conviction that the future relied on science and technology, many scientists and engineers may have viewed the onset of this current economic storm with some complacency.

Well, no more. Unlike previous recessions, this one has reached into academia directly through budget cuts to higher-education institutions, especially in the United States at, for example, state-funded universities. And the negative effect on research funding may persist longer than normal. Despite efforts in the United States to boost science funding via stimulus spending, federal and state governments will be left with historically large budget deficits to manage: the pressure to cut discretionary government spending will be enormous. And with big losses to retirement portfolios, many senior scientists and engineers will probably remain at their benches for a few years longer to rebuild their nest eggs.

Yes, the news is grim. But I would argue that this is one of the most important events for science in the past 50 years. Scientists and engineers are coming to realize that the traditional models of a 'scientific' career are outdated. Science careers in the future will be marked by adaptability, entrepreneurialism and self-reliance. It's more than a realignment to a different set of jobs; it's a change in approach to a career. The era of 'career resilience' has begun.

## From I-shaped to T-shaped

When my father embarked on his scientific career, he was advised to choose one topic and aspire to be the best in the world at it. This made great sense. Science was growing in all directions, so even an infinitesimally narrow area of knowledge was bound to expand. The prevailing wisdom was to be an 'I'-shaped scientist, with deep expertise in a single area.

Today, focusing on a single area of science no longer guarantees safety and security. Like an investment portfolio made up of just one stock, careers that are I-shaped are

prone to large swings in value. Career experts often recommend that professionals adopt a 'T'-shaped strategy: deep technical expertise is still needed, but capping that depth should be a broad set of interdisciplinary skills that allow a scientist or engineer to solve problems in a wide range of applications. Despite the growth of interdisciplinary centres on some campuses, university training has yet to adapt to this new reality. Traditional PhD programmes still train students to focus on a single area of scholarly research. 'Broadening' activities, such as participation in interdisciplinary studies or taking business or law courses is, at best, tolerated.

## How to cap your 'T'

Broadening your experience base and your network is now widely recognized to be the most effective way to expand your career options and make your career more resilient.

Fortunately, there are several ways to approach this that don't need a major time investment.

Technical consulting, for instance, expands your network and skill base while allowing you to apply your experience and problem-solving skills to real-world problems. Many scientists think that their area of expertise does not have any commercial value or relevance. In a narrow sense, that may be true. But scientists and engineers often fail to appreciate the breadth of expertise that they have developed. Interpreting analytical data, formulating and testing mathematical models, even literature review and analysis, are activities that many private-sector consultants engage in (and are well-compensated for).

Uncertain how to proceed? Seek out someone in your department, school or field who has an active consulting practice and find out how they discover opportunities.

Time away from the lab is often key. Rotational assignments such as sabbaticals provide academics with an opportunity to work in a different institution, build new skills and expand networks. Today, many academics use their sabbatical simply to escape their teaching duties for a year and never leave their department. Even short periods spent at another institution, university or company can be enormously broadening, both for your skills and for your network. Exchange assignments as short as four weeks can allow you to meet and interact with numerous new colleagues.

Scientists and engineers tend to be somewhat reserved in building their networks — focusing mainly on whom they know in the research community. In fact, it is the people outside the world of research who may have greatest effect on your career. If you are interested in an area of technology or commerce outside your area of expertise, consider attending a technical meeting on the subject. Interested in the interface of science and business? Attend lectures at the nearest business school. You'd be surprised who you will meet just by mingling.

Career resilience means investing in a range of activities. It means cultivating and maintaining an active network of colleagues, partners, friends and supporters who can steer you towards new opportunities. It means thinking of yourself as an adaptable and clever problem-solver rather than a technical expert with a narrow sliver of knowledge.

**Peter Fiske is chief technology officer of PAX Water Technologies in San Rafael, California, and author of *Put Your Science to WORK*.**



## NEWS

# Networking in VIVO

Friends follow each other on Facebook. Colleagues swap contact details on LinkedIn. And soon, academic scientists eager to forge cross-disciplinary collaborations will have their own network: VIVO.

Funded by a US\$12.2-million grant from the US National Center for Research Resources, VIVO aims to allow researchers to transcend disciplines and tap into the collective resources needed to facilitate breakthrough discoveries. The site was first created at Cornell University in Ithaca, New York, in 2003, but for use only at the university ([vivo.cornell.edu](http://vivo.cornell.edu)). The money, allocated through the American Recovery and Reinvestment Act of 2009, will be used to expand it across universities and to improve its cross-disciplinary search capabilities.

Its creators say that VIVO will help to create the collaborations that are increasingly crucial in science. Michael Conlon, lead investigator on the grant and interim director of biomedical informatics at the University of Florida in Gainesville, argues that web searches generally require too much sifting through results to find the specific researcher capabilities desired for a given project. A VIVO search allows researchers to quickly search for faculty members who have a specific expertise or skill set — for example, a genomics specialist looking to team up with a bioinformatics specialist and a clinician.

Unlike Facebook or LinkedIn, VIVO-

assembled web pages will be filled with information from official, verifiable sources. For example, academic institutions will provide information about researchers' positions, and

publication lists will be pulled directly from scientific journals. This should produce results with much greater specificity than, say, using Google to mine web pages, says Conlon. The site's algorithm searches for scientific topics, and VIVO yields detailed 'hits' broken down into the most relevant subcategories — such as

researchers, activities, events, organizations, publications or subtopics related to the term. "If you search for people, you'll just get people," says Conlon.

Researchers will be able to individualize their sites with detailed information about interests and research opportunities. VIVO membership criteria are still being discussed, and the requirements may differ between participating institutions, but all academics and graduate students should be eligible.

Conlon says that by next year, VIVO will connect seven universities in California, Florida, Indiana, Missouri, New York and Puerto Rico. He hopes for many more by the end of the two-year grant, including non-US schools.

Although Conlon cringes at the notion that VIVO will become a frivolous 'Facebook for scientists', he admits that they will use Facebook to spread the word. ■

Virginia Gewin



IMAGES.COM/CORBIS/B. REA

## POSTDOC JOURNAL

# The career less travelled

In Robert Frost's *The Road Not Taken*, a traveller reaches a fork in a wooded path, and chooses the less-worn trail. Walking along, he imagines a future in which he will say: "I took the one less travelled by/ And that has made all the difference." The poem seems to say that independence leads to happiness.

Yet on closer reading, this interpretation seems unlikely. Far from one path being less travelled, both are worn "about the same" after the traveller's passing. The path becomes less travelled only in the traveller's imagination, as

he dreams of a moment "ages and ages hence", implying that independent thought leads to happiness. To me, the poem suggests that we often ascribe meaning and purpose to arbitrary decisions.

At a conference recently, I heard a graduate student ask a professor to describe the choices that led to his career success. I smiled when the professor replied that many of his key decisions were determined by chance. The student seemed surprised, but to me it made sense. Unlike Frost's traveller, the professor was honest about

the role of chance in his life.

As I begin a new year of faculty applications, I must write a research plan. In this short essay, I have to describe the course of my past research as well as of future work. Reflecting on past decisions relating to my research direction, I find myself in the position of Frost's traveller, trying to ascribe meaning and purpose to the machinations of fate. My career path owes an awful lot to chance. ■

**Sam Walcott** is a postdoc in theoretical biophysics at Johns Hopkins University in Baltimore, Maryland.



## IN BRIEF

### Academic benefits

Three universities and a government research organization are in the top 10 of the 'best US employers' list for workers aged older than 50, as compiled by the AARP, formerly the American Association of Retired Persons. Cornell University in Ithaca, New York, came first, the US National Institutes of Health third, Massachusetts Institute of Technology seventh and George Mason University of Fairfax, Virginia, tenth. GlaxoSmithKline at Research Triangle Park, North Carolina, debuted this year at 25. The annual survey, launched in 2001, cited retirement benefits such as health and life insurance, job training and placement for older workers. The list is at [go.nature.com/h42oIj](http://go.nature.com/h42oIj).

### Topping out

The retention of US science, technology, engineering and maths (STEM) students is strong except among the best performers, says a study that followed student cohorts from high-school graduation as early as the 1970s through to mid-career as late as 2005. The report, called *Steady as She Goes?*, debunks the notion that the United States is not producing enough STEM college graduates, say its authors. But it suggests that many high performers left in early to mid-career in the late 1990s. The authors, Hal Salzman of Rutgers University in New Jersey and Lindsay Lowell of Georgetown University in Washington DC, speculate that they may have opted for non-traditional science careers or higher-paying jobs.

### Changes planned for ERC

An independent review of the nascent European Research Council (ERC) — launched in February 2007 — has prompted the European Commission to change the council's management and regulatory structure. Its executive agency had been led by a secretary-general, but it is now seeking a scientist with leadership experience for the top post to combine scientific and administrative functions, according to the commission. The council will also seek input from the European Parliament and the European Council of Ministers on establishing regulations for funding research. The council's funding budget, provided by the commission, was set up under a European Union programme providing €7.5 billion (US\$11 billion) for the period 2007–13.



# Clear proof

The final demonstration of the failure of cold fusion.

**Jeff Hecht**

Repetition of experiments is important both to verify results and to resolve any questions about their interpretation. This is particularly vital for the experiment shown in the most-watched video on the Internet, 'The Final Demonstration of the Failure of Cold Fusion', as expounded by Professor Madeline Hou.

Recorded in standard-definition two-dimensional colour, the video doubtless benefits from Professor Hou's presentation. A born entertainer with a gift for showmanship, at the time of the recording she was strikingly attractive with jet-black hair that hung down almost to her waist. But the real allure is the demonstration itself.

"No experiment is perfectly repeatable," she begins. "On a fundamental level, the uncertainty principle says we cannot simultaneously measure position and momentum of any particle with infinite precision. On a practical level, it is impossible to control all variables that might conceivably affect experimental results, such as the weather here in Boston in January."

Although the video does not show the students in the lecture hall, it picks up their laughter. National Weather Service records show that six inches of snow fell that morning, following a prediction of 'partly cloudy'.

"This is the eighth in my series of annual demonstrations of the failure of cold fusion," she says, flirting with the camera. "Although I use the same equipment, the same procedures, and even wear the same clothing, it is impossible to control every possible influence, such as the flux of wireless signals from your laptops, smartphones and personal contact interfaces."

She unveils a classic Pons–Fleischmann cold-fusion cell, displaying heavy water, palladium electrodes and a calorimeter linked to a large classroom meter with an old-fashioned analogue galvanometer needle. "We know that this equipment did not produce heat beyond that expected from classical electrochemistry in my seven past Intersession demonstrations. Yet Pons and Fleischmann claimed to observe the release of excess energy, which they attributed to the fusion of two deuterium atoms to yield tritium and a neutron. That reaction should have released a dangerous flux of neutrons. Where were the dead graduate students?"

Chuckles are audible from the mostly undergraduate audience.

Hou smiles. "We need to consider the possibilities. Other experimenters claim to have observed excess energy, but curiously they cannot reproduce their own experiments reliably. Could they have failed to note special conditions needed to generate excess heat? What might such conditions be?"

It was a rhetorical question, but she could count on the students to shout out answers. "True belief in cold fusion," says one. "Freedom from doubt," says another.



"The absence of verification," says a third. Comparison with recordings of prior lectures shows these are the first, second and third most common answers over the eight-year series. Students who attended the seventh demonstration say that chanting answers in that sequence was a campus ritual.

Hou also is carefully choreographed. She shows video clips of the palladium rods being taken from the previous year's apparatus, heated in a vacuum to remove the deuterium, reloaded with the gas and mounted in the electrochemical apparatus. Verified digital date stamps on each clip show when the procedures were done.

"We are now prepared for the moment of truth," she says. "Of course, you must realize that the precise value of truth in experimental practice is subject to statistical variation, but after an experiment has been conducted several times, it is possible to rule out some values with reasonable confidence. What observations can we rule out if the experiment is properly conducted?"

The students shout out a cacophony of answers. "Melting the apparatus", "cooling the room to one nanokelvin", "nuclear fusion in the air" and "nothing" are audible.

"Personally, I continue to believe that there is a small but finite probability that variations in results from repetition of the experiment will teach us something, although that probability declines with each repetition," Hou says. "However, we can be confident that it will not release as much energy as the Hiroshima bomb."

As the audience chuckles, the camera zooms in to focus on Hou's right hand poised on the handle of a large antique knife switch inscribed 'Tesla Co.'. She pulls the switch down to apply a voltage across the cell.

The video switches to a split screen, with one side showing the meter and a timer, and the other the cold-fusion cell. For the first 42 seconds, neither image changes. At the 43rd second, the needle moves slightly above zero and a small bubble appears in the cold-fusion cell. More bubbles appear in succeeding frames, and the needle moves upwards.

"That's funny," Hou's voice says, off-camera. A smartphone QuickVid sent in real time by a student shows a surprised look spread across her face as she looks at the moving meter. The bubbles multiply and the meter accelerates. "Uh-oh," Hou says, reaching towards the switch.

A blinding flash ends the video and smartphone transmissions.

Substantial uncertainty remains in the estimates of the energy released by the Final Demonstration, although measurements of the crater indicate that it was significantly less than the Hiroshima bomb. Unfortunately, the Final Demonstration vaporized all the experimental apparatus, and no records survive of activities being conducted in the building or of wireless signals in the vicinity. Nonetheless, we can be confident that the energy release was orders of magnitude higher than the total from any previous 'controlled fusion' experiment.

With new energy sources critically important, the Department of Energy has scheduled its attempt to replicate Hou's demonstration at the Nevada Nuclear Test Site.

**Jeff Hecht is a science and technology writer based in the Boston area.**

**Join the discussion of Futures in Nature at [go.nature.com/QMAM2a](http://go.nature.com/QMAM2a)**

JACEY